THE ACCOUNTING REVIEW

VOLUME 88, No. 2 MARCH 2013

A JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice.

Membership in the Association is available at the following annual rates.

Full membership:

Membership dues of $205.00 include *Accounting Education News*, and the online selection of all three Association-wide journals: *The Accounting Review*, *Accounting Horizons*, and *Issues in Accounting Education*. (This does not include Section journals.) Full members may also opt to receive *hardcopies* of AAA journals for an additional cost:

The Accounting Review	with Print Option $45.00	Total U.S. $250.00
Accounting Horizons	with Print Option $25.00	Total U.S. $230.00
Issues in Accounting Education	with Print Option $25.00	Total U.S. $230.00
All three Association-wide journals	with Print Option $95.00	Total U.S. $300.00

Associate membership:

Full-time students are eligible for associate membership. The print option is not available for Associate members.
All three Association-wide journals . . . . U.S. $35.00

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Auditing; Diversity; Financial Accounting and Reporting; Forensic and Investigative Accounting; Gender Issues and Worklife Balance; Government and Nonprofit; Information Systems, International Accounting; Management Accounting; Strategic and Emerging Technologies; Teaching, Learning and Curriculum; and Two-Year College. Each Section assesses its own dues, which are collected by the Association.

Editorial correspondence and revisions of manuscripts originally submitted before June 1, 2011 should be sent via email to the incoming Senior Editor, John Harry Evans III, University of Pittsburgh, at tar@katz.pitt.edu • Telephone: (412) 624-8605 • Fax: (412) 648-1693. New manuscripts must be submitted through the Manuscript Submission and Peer Review System, located at http://accr.allentrack.net. The site contains detailed instructions regarding the preparation of files for submission.

The submission fee of $200.00 for members of the AAA or $400.00 for nonmembers may be paid online (VISA or MasterCard only) at https://aaahq.org/AAAforms/journals/tarsubmit.cfm and should be paid before the article is submitted.

Requests for information concerning membership, subscriptions, dues, advertisements, and other matters related to the Association (other than submission of manuscripts) should be sent to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Info@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review*.

The Accounting Review is indexed in the Social Science Citation Index, which is included in Web of Science.

The Accounting Review (ISSN 0001-4826 print and ISSN 1558-7967 online) is published six times a year in January, March, May, July, September, and November by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rate of $430 for a print copy. The member subscription rate for a hard copy is $45. Electronic access is included with membership dues. Periodicals Postage Paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address changes to ***The Accounting Review***, 5717 Bessie Drive, Sarasota, Florida 34233-2399.

THE ACCOUNTING REVIEW

A JOURNAL OF THE

AMERICAN ACCOUNTING ASSOCIATION

VOL. 88 MARCH 2013 No. 2

PAST EDITORS

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 365–393

American Accounting Association
DOI: 10.2308/accr-50347

Insider Trading, Litigation Concerns, and Auditor Going-Concern Opinions

Chen Chen
University of Auckland

Xiumin Martin
Washington University in St. Louis

Xin Wang
The University of Hong Kong

ABSTRACT: We investigate whether insider selling affects the likelihood of firms receiving auditor going-concern opinions. Prior studies document significant negative market reactions to the issuance of going-concern opinions, indicating that such opinions convey bad news to investors. Insider sales followed by negative news are likely to attract regulators' scrutiny and investor class-action lawsuits. Therefore, we predict that, to reduce the risk of litigation, managers have incentives to avoid receiving going-concern opinions after their insider sales by pressuring auditors for clean audit opinions. We evaluate this prediction empirically and find that the probability of receiving a going-concern opinion is negatively associated with the level of insider selling. Further analysis indicates that this negative relation is more pronounced for firms that are economically significant to their auditors but less pronounced when (1) auditors have concerns about litigation exposure and reputation loss and (2) audit committees are more independent. Finally, the negative relation between going-concern opinions and insider sales is significantly weakened after SOX.

Keywords: *insider trading; litigation risk; going-concern opinion; SOX.*

JEL Classifications: *G18; M42; G48.*

I. INTRODUCTION

We investigate whether managers' litigation concerns about insider selling affect the likelihood of firms receiving going-concern opinions. Prior studies show that managers face the risk of trade-related litigation around news events (Seyhun 1992; Givoly and

---

We thank David Hay, Steven Cahan, Ron King, Rich Frankel, Srinivasan Sankaraguruswamy, and workshop participants at National University of Singapore, The Chinese University of Hong Kong and The University of Auckland for their helpful comments. Professor Wang acknowledges the financial support from The University of Hong Kong.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: June 2010
Accepted: September 2012
Published Online: October 2012

Palmon 1985). To reduce their risk exposure, managers have at least two options. First, they can abstain from trading before notable events. Seyhun (1992), for example, finds that insiders reduce their "timely" trades before major events, such as earnings and takeover announcements (i.e., sales before negative news and purchases before positive news). Alternatively, when managers do choose to trade, they can attempt to alter the information flow in the post-trading period to avoid price swings and escape regulators' scrutiny.[1] We focus on this option with respect to the association between managers' insider sales and auditors' going-concern modifications.

Prior studies (e.g., Firth 1978; Chow and Rice 1982; Fleak and Wilson 1994; Jones 1996; Carlson et al. 1998; Menon and Williams 2010) provide evidence of negative market reaction to first-time going-concern reports.[2] Researchers have also documented that managers can influence auditors' opinions and engage in opinion shopping (Carcello and Neal 2000; Lennox 2000). Given that auditors tend to make Type I errors—studies show that a low percentage of firms receiving going-concern reports declare bankruptcy in the year following the audit opinions (e.g., Myers et al. 2011)—opinion shopping seems particularly plausible when the nature of the opinion is ambiguous. If managers have incentives to avoid insider-selling related litigation, then we expect that they will exert a greater influence on auditors to avoid receiving a going-concern report. In other words, we predict an inverse relation between the likelihood of receiving a first-time going-concern opinion and insider selling. Anecdotal evidence seems to support our prediction. For example, when companies file registration statements so that insiders can sell stock, they often cite clean audit reports as supporting documents.[3]

At least two reasons motivate our focus on insider selling. First, the information content of the two types of auditor opinion is asymmetric. First-time going-concern opinions induce significantly negative market reactions, while clean opinions do not generate positive market reactions (Firth 1978; Jones 1996; Menon and Williams 2010). Insiders are therefore less concerned about buying and the subsequent receipt of a clean opinion. Second, Roulstone (2008) argues that bad-news disclosures are more likely to trigger investor lawsuits that allege inadequate disclosure by management. Such lawsuits usually use pre-disclosure insider selling to indicate management's foreknowledge of bad news. Thus, in contrast to insider purchases ahead of good news, insider sales ahead of bad news carry a significant legal risk.[4]

Using a sample of 12,329 firm-year observations based on a two-step regression technique, we find evidence that a higher level of insider selling is associated with a lower likelihood of receiving a first-time going-concern report. For a one standard deviation increase in insider selling, the probability of receiving going-concern reports decreases by 1.39 percent. We perform four

---

[1] Regulators frequently detect insider trading by examining the price movement of a company's stock and identifying suspicious events, and shareholders are likely to file class action lawsuits following a drastic decline in stock price.

[2] Anecdotal evidence is consistent with these findings. For example: "Drkoop.com filed its Annual Report . . . that its auditors doubted the Company's ability to continue as 'going concern.' Following these revelations, the company's stock plummeted from a previous close of \$6.25 to a close of \$3 on March 31, 2000—a one day drop of approximately 41 percent" (*Business Wire* 2000).

[3] When Quepasa.com filed a new registration statement in September 2000 to enable some insiders to sell stock, KPMG, the *fourth* auditor within a year, allowed the company to cite its clean audit opinion from early 2000. Quepasa.com was closed by the end of 2000 (Weil 2001; see Quepasa.com's filing of form S-3 on Sept. 15, 2000). In fact, it is fairly common that insiders incorporate auditors' reports into their filings with the SEC to disclose their intent to sell. For example, we find such cases for Argon St. Inc in 2005 and Vector Group Ltd in 2004.

[4] A law firm (Wolf Haldenstein Adler Freeman & Herz LLP) filed a class action lawsuit against Drkoop.com Inc., alleging that the company and certain insiders had provided materially false and misleading information before the firm received a going-concern opinion. Some of the supporting evidence was that insiders liquidated their stock positions in the company immediately preceding the issuance of a going-concern opinion (*Business Wire* 2000).

cross-sectional analyses. The results indicate that the negative relation between insider selling and the probability of receiving a going-concern opinion is stronger for firms that are more economically important to their auditors but weaker for firms whose auditors have greater concerns about litigation exposure and reputation loss and for firms with more independent audit committees. Next, to address the direction of causality, we examine auditor switches following the issuance of clean opinions. Specifically, we attempt to determine whether managers exert influence over auditors' opinions to reduce their risk of selling-induced litigation or insiders to reduce their selling in anticipation of a going-concern opinion. We find that auditors who issue clean opinions for clients with higher levels of insider selling have a lower frequency of dismissals in the subsequent year. These results are consistent with the notion that management influences auditors' opinions but are inconsistent with the notion that insiders reduce their selling in anticipation of going-concern reports.

We also find that the negative relation between insider selling and the likelihood of receiving a going-concern opinion holds for the pre- and post-SOX periods but is significantly weaker in the post-SOX period. We attribute this finding to the overall strengthening of corporate governance, which enhances the independence of both the internal audit committee and the external auditors (Chambers and Payne 2008). To alleviate the concern that firms receiving going-concern opinions are fundamentally different from those receiving clean opinions, we re-run our main analysis based on an alternative matched sample and the marginal effect of insider trading on going-concern reports are even stronger. In additional analyses, we find more pronounced results for insider trades occurring in the later part of the year and for relatively healthier firms. The latter finding supports the notion that, when pressured by a client, an auditor might be lenient if the situation giving rise to a potential going-concern report is ambiguous. Finally, our main findings are robust to an estimation based on a simultaneous equation system, to an estimation controlling for earnings management and the management's plans, and to the use of alternative measures of insider trading.

Two caveats are in order before proceeding. First, we offer no direct evidence that managers pressure their auditors when engaged in insider sales. Although the results suggest indirectly the existence of such pressure, our findings are also consistent with an alternative explanation. That is, when observing insider sales prior to issuing an opinion, auditors may decide to be less conservative in their opinion by rationally trading off the expected cost of litigation associated with insider sales preceding a going-concern opinion with that of a clean opinion for a firm going bankrupt. Alternatively, corporate insiders may simply trade for liquidity reasons that lead to a higher level of insider selling for firms with clean opinions than for going-concern firms. At the same time, these clean firms have better fundamentals and, thus, we observe a lower likelihood of receiving a going-concern opinion. Second, our two-step regression technique, which intends to address reverse-causality concerns, has the downside that the economic factors used to predict net insider sales are largely legitimate and observable drivers of insider selling rather than reflecting insiders' opportunistic behavior. We attempt to mitigate this concern and gauge insiders' opportunistic trading behavior by using changes in lieu of levels to calculate net insider sales. The limitation, however, cannot be completely eliminated due to the nature of the two-step regression technique.

Our study offers two primary contributions. First, it provides insight into the incentive effect of corporate insider trading on auditor behavior. Prior studies examine auditor behavior from the perspective of the corporate governance of auditees (Carcello and Neal 2000) and auditors' economic incentives (DeFond et al. 2002). However, there is little evidence on the effect of managers' incentives on auditor behavior. Our study helps to fill this gap by providing evidence on the relationship between managers' incentives and auditors' opinions.

Second, this study adds to the literature on insider trading. Prior studies show that insiders' incentives to maximize trading profits affect earnings management (Beneish and Vargus 2002;

Beneish et al. 2011) and voluntary disclosure (Cheng and Lo 2006).[5] Rogers (2008) documents that managers make higher quality disclosures before they sell shares to reduce litigation risk. A closely related study by Stanley et al. (2009) examines insider trading surrounding first-time going-concern opinions and finds that insider sales immediately before and after going-concern opinions predict bankruptcy. Our evidence extends this literature by showing that insiders' incentives to sell and their desire to avoid litigation can influence auditors' reports.

Section II develops our hypotheses. Section III describes our sample and research method. Section IV reports the results. Section V presents additional analyses, sensitivity tests, and further discussion. Section VI concludes.

II. HYPOTHESIS DEVELOPMENT

Managers can capitalize on their private information through insider trading. Prior research suggests that managers time stock trades by buying before price increases and selling before price declines.[6] However, insider trading regulations impose penalties for illegal insider trades. For example, the Insider Trading and Securities Fraud Enforcement Act of 1988 increased the penalties for illegal insider trading by 300 percent and increased criminal penalties by 1,000 percent (Jagolinzer and Roulstone 2009). Seyhun (1992) and Garfinkel (1997) examine the association between regulatory enforcement and insider trading volumes before earnings releases and find evidence of decreased "timely" trades in anticipation of earnings news. Insiders can avoid regulatory sanction by abstaining from trading. But insiders can also coordinate timely trades and the release of news during the post-trading period to achieve the same goal. In other words, by altering the post-trade information flow, insiders can attempt to avoid signals that regulators or investors use to detect insider trading.

DeFond et al. (2002) note that auditors' opinions play an important role in warning market participants of impending going-concern problems. Prior studies show that stock prices respond negatively to the issuance of first-time going-concern reports (Firth 1978; Fleak and Wilson 1994; Jones 1996; Menon and Williams 2010). Fleak and Wilson (1994) argue that auditors' going-concern opinions can affect stock prices for at least two reasons. First, a going-concern qualification provides a warning about the auditor's assessment of a company's probable viability. Second, a going-concern opinion can have direct negative consequences for a firm's future cash flow, resulting in technical default on existing debts or a defective SEC registration statement. Consistent with this argument, Menon and Williams (2010) find −3.49 percent one-day (+1) and −6.28 percent three-day (0, +2) size-adjusted stock returns following the release of going-concern reports. Compared with the criteria of −20 percent quarterly abnormal returns for large price drops used in Beck and Bhagat (1997), who study shareholder litigation, a three-day abnormal return of −6.28 percent is nontrivial. Several other studies show similar results. Firth (1978) finds that announcements of going-concern opinions induce an average abnormal return of −4.1 percent for 35 firms from the U.K. Jones (1996) documents an average abnormal return of −2.79 percent over (−2, +2) days around going-concern opinion announcements for 68 U.S. firms.

---

[5] Another extensive literature argues that insider trading increases litigation risk and that insiders attempt to avoid trading around significant information events, including takeover announcements (Seyhun 1992), bankruptcies (Seyhun and Bradley 1997), management forecasts of earnings (Penman 1985; Noe 1999), earnings announcements (Park et al. 1995; Garfinkel 1997), and breaks in positive earnings strings (Huddart et al. 2003).

[6] These studies include Jaffe (1974), Finnerty (1976), Baesel and Stein (1979), Givoly and Palmon (1985), Seyhun (1986), Rozeff and Zaman (1988), Seyhun (1998), Beneish and Vargus (2002), and Huddart et al. (2003), among others.

Carcello and Neal (2000) and Lennox (2000) find that auditors are less likely to issue going-concern opinions when pressured by their client firms. Furthermore, given the low percentage of firms that file for bankruptcy after receiving a going-concern report (Stanley et al. 2009; Myers et al. 2011), auditors tend to err on the conservative side and make Type I errors. Thus, auditors likely have some discretion in issuing going-concern reports, particularly when the nature of the opinion is ambiguous. If managers have incentives to avoid insider-selling-related litigation, then we expect that they will pressure auditors for clean opinions. This line of argument leads to the first hypothesis, formally stated as follows:

H1: Insider selling activity is negatively associated with the likelihood of receiving a going-concern opinion.

Auditors trade off the benefits, such as client retention, and costs, such as litigation and reputational damage, when deciding whether to issue a going-concern report (DeAngelo 1981; Watts and Zimmerman 1981). As a consequence, the negative relation between insider selling and the likelihood of receiving a going-concern opinion is expected to be more pronounced when clients are economically important.

On the other hand, high litigation costs and reputation concerns are likely to motivate auditors to act conservatively (Antle et al. 1997). Shu (2000) finds that auditors resign from clients in response to increased litigation risk and emerging mismatches with the clients. Based on these arguments, we conjecture that the negative relation between insider selling and the probability of receiving a going-concern opinion is likely to be attenuated when auditors have greater concerns about litigation costs and reputational harm. In summary, the above arguments motivate the next three hypotheses.

H2: The inverse relation between insider selling and the probability of receiving a going-concern opinion is more pronounced for firms that are economically more important to their auditors.

H3: The inverse relation between insider selling activity and the probability of receiving a going-concern opinion is less pronounced for firms whose auditors have higher litigation costs.

H4: The inverse relation between insider selling activity and the probability of receiving a going-concern opinion is less pronounced for firms whose auditors have better reputations.

An important role of audit committees is to protect external auditors from dismissal following the issuance of an unfavorable report. Carcello and Neal (2000) find that audit firms are less likely to issue going-concern reports to financially distressed clients whose audit committees lack independence. Furthermore, Carcello and Neal (2003) find that audit committees with greater independence are more effective in shielding auditors from dismissal after the issuance of new going-concern reports. Taken together, their evidence suggests that auditors are less likely to be influenced by managers when the audit committee is more independent. SOX Section 301, which requires public companies to have a fully independent audit committee, reflects regulators' belief that an independent audit committee contributes to higher audit quality. Following this line of argument, our fifth hypothesis is formally stated as follows:

H5: The inverse relation between insider selling and the probability of receiving a going-concern opinion is less pronounced for firms whose audit committees are more independent.

III. SAMPLE SELECTION AND RESEARCH METHOD

Sample Selection

We obtain insider trading data from Thomson Reuters and examine two types of insider trading transactions in our analysis: open market purchases and open market sales.[7] To investigate the aggregate influence of management, we sum the purchases and sales by all top managers of the same firm over a fiscal year. Following prior studies (e.g., Cheng and Lo 2006; Huddart et al. 2007), we use the dollar amount of insider trades to capture the strength of trading incentives.

We obtain information about audit opinions and audit fees from Audit Analytics for the period 2000 through 2007. We then match the audit opinion data with the Compustat industrial annual file for the necessary financial statement variables, the Center for Research in Security Prices (CRSP) database for stock return variables, and the Insider Trading database for insider trading activity. This procedure leaves a total of 35,188 firm-year observations with the necessary variables. As in prior research, we restrict our sample to financially distressed firms because the decision to issue a going-concern opinion is most salient among these firms (Hopwood et al. 1994; Mutchler et al. 1997; Reynolds and Francis 2000; Carcello and Neal 2000; DeFond et al. 2002).[8] As in DeFond et al. (2002), we classify a firm as financially distressed when it reports either negative earnings or negative operating cash flow during the current fiscal year. Firm-years following first-time going-concern opinions are excluded because our focus is on auditors' decisions to issue first-time going-concern opinions. These restrictions exclude 21,489 firm years that are not financially distressed and 1,320 firm years following first-time going-concern opinions. Our final sample retains 12,329 firm years, consisting of 801 observations with first-time going-concern opinions and 11,528 observations with clean opinions.

Research Design

H1 predicts an inverse relation between the likelihood of receiving a going-concern opinion and insider selling activity. We measure insider trading by aggregating the trades of top managers over one year, ending with the current year's earnings announcement.[9] Because courts use so-called abnormal insider sales to establish scienter in securities lawsuits, we gauge the abnormal portion of insider sales (*CHNSV*) by the difference between the net selling volume (*NSV*) of the current year and the average *NSV* during the prior four years. The net selling volume (*NSV*) is the natural logarithm of 1 plus the absolute total dollar amount of net insider trading (i.e., insider purchases subtracted from insider sales, in thousands of dollars) over a fiscal year, with a positive (negative) sign added for net sales (net purchases).

---

[7] We include insider trading transactions by top officers—that is, chairman, vice chairman, CEO, COO, president, and CFO—but exclude transactions by lower-tier officers and non-officer insiders who are unlikely to influence management's private incentives. Chairman, vice chairman, CEO, COO, and president are classified as Level I insiders in the insider trading database. Our results do not change when we use CEOs' insider trades only or when we use the aggregate insider trades by all corporate insiders.

[8] Financially distressed firms may have less bargaining power with auditors over an audit opinion. Consequently, the sample consisting of these firms may be less appropriate for examining our research question. However, given that Carcello and Neal (2000) demonstrate the effect of audit committee independence on the likelihood of auditors issuing going-concern opinions among financially distressed firms, we are less concerned about this issue.

[9] Auditors commonly determine an audit opinion in two to three months after a fiscal year-end, and their final opinion is usually filed after earnings announcements. We therefore treat the after-year-end trades that occur before earnings announcement as part of the current year's insider activities because these trades may still influence auditors. Untabulated analysis shows that our results are robust to the exclusion of this part of trades.

To test the effect of insider trading on the likelihood of receiving a going-concern opinion, we must control for trading activity that occurs in anticipation of an auditor's opinion (i.e., reverse causality). For example, the fact that a higher level of insider sales in a fiscal year is correlated with a lower likelihood of receiving a going-concern opinion may simply indicate either one or both of the following: (1) managers who want to sell discourage auditors from issuing going-concern opinions, as hypothesized in H1, or (2) when managers anticipate that an auditor is likely to issue a going-concern opinion, they are less likely to dispose of their holdings. We address this endogeneity issue by using instrumental variables (IVs) for our insider trading variable. Specifically, we employ a two-step regression technique following Cheng and Lo (2006). In the first step, we obtain the predicted change in insider sales (*PREDICT_CHNSV*) for the current fiscal year, using information from prior fiscal years such as option grants, firm size, growth opportunities, and stock returns. The model specifications, variable definitions, and results are presented in Appendix A. This procedure results in valid IVs for insider trading, as the predicted change in insider sales in year t is correlated with the actual change in net insider selling volume (*CHNSV*) in year t, and it is not endogenously affected by the type of audit opinion issued for year t. In the second step, we follow DeFond et al. (2002) and estimate the probability of receiving a going-concern opinion using the *predicted* change in insider selling (*PREDICT_CHNSV*) from the first-stage regression for financially distressed firms:

$$\begin{aligned}Pr(GC\ OPINION_t = 1) = {} & \beta_0 + \beta_1(PREDICT\_CHNSV_t) + \beta_2(ZSCORE_t) + \beta_3(LOSS_t) \\ & + \beta_4(SIZE_t) + \beta_5(AGE_t) + \beta_6(RETURN_t) + \beta_7(VOLATILITY_t) \\ & + \beta_8(LEV_t) + \beta_9(CLEV_t) + \beta_{10}(OCF_t) + \beta_{11}(ANNLAG_t) \\ & + \beta_{12}(INVESTMENTS_t) + \beta_{13}(NEWFINANCE_t) + \beta_{14}(BIGN_t) \\ & + \sum Industries + \sum Years + \varepsilon_t, \qquad (1)\end{aligned}$$

where:

GC OPINION = an indicator variable equal to 1 for firms receiving a first-time going-concern audit opinion, and 0 otherwise;

PREDICT_CHNSV = predicted value of the change in net insider selling volume (*CHNSV*) based on the model of insider trades in Appendix A;

ZSCORE = Altman's Z-score (Altman 1968) for the current year;

LOSS = an indicator variable coded 1 if a firm reports a negative net income (Compustat item NI) for the current year, and 0 otherwise;

SIZE = natural logarithm of total assets at the end of the current year (Compustat item AT);

AGE = natural logarithm of firm age, which is proxied by the number of years of data in Compustat from the start of coverage to the current year;

RETURN = firm's cumulative stock return over the current year;

VOLATILITY = standard deviation of monthly returns over the current year;

LEV = ratio of total liabilities (Compustat item LT) to total assets at the end of the current year;

CLEV = change in *LEV* from the previous year to the current year;

OCF = operating cash flow (Compustat item OANCF) divided by total assets for the current year;

ANNLAG = number of days between the fiscal year-end and earnings announcement date for the current year;

INVESTMENTS = short- and long-term investment securities (including cash and cash equivalents) (Compustat items CHE and IVPT), scaled by total assets;

NEWFINANCE = an indicator variable equal to 1 if a client has a new issuance of equity or debt in the subsequent fiscal year (i.e., nonzero Compustat item DLTIS or the amount of

Compustat item SSTK exceeding 5 percent of the firm's market value of equity), and 0 otherwise; and

BIGN = an indicator variable equal to 1 if the auditor is a member of the Big 5 before 2002 or a member of the Big 4 after 2002, and 0 otherwise.

Equation (1) is estimated using a pooled logistic regression, and the significance level of the coefficients is derived based on robust standard errors clustered by the client firm and the auditor. Appendix B provides further details to motivate the explanatory variables.

IV. RESULTS

Descriptive Statistics

Table 1, Panel A presents descriptive statistics for the distressed firm sample used in estimating Equation (1). We winsorize all continuous variables at the 1st percentile of each side of the statistical distribution, except for *RETURN*.[10] Our sample has a positive mean of 0.167 for the predicted change in net insider selling (*PREDICT_CHNSV*), where *PREDICT_CHNSV* is derived from estimating Equation (A1) in Appendix A. The change in net insider selling (*CHNSV*) has a mean value of –$133,560, and the mean value for the log-transformed is –0.214.[11] About 74 percent of the firm-years have Big N auditors, and 87 percent of firm-years report a loss. Panel A also shows that 58.1 percent of the distressed firms have new financing activities in the subsequent year. The statistics of the other variables in Equation (1) are comparable to those reported in DeFond et al. (2002).

Table 1, Panel B reports descriptive statistics by auditor opinion. In total, 11,528 firm-years receive a clean audit opinion (hereafter, clean firms) and 801 firm years receive a going-concern audit opinion (hereafter, GC firms). Panel B shows that GC firms and clean firms display significant differences in all of the variables used in the regression. Specifically, GC firms have a lower predicted change in net selling activities (*PREDICT_CHNSV*) than clean firms (–0.131 versus 0.189).[12] This evidence is consistent with our prediction that the probability of receiving a going-concern report decreases in insider sales. We also observe a lower level of insider selling for GC firms than for clean firms using the actual change in net insider selling (*CHNSV*) (–0.505 versus –0.191 for the log value and –$319, 590 versus –$114, 250 for the raw value). The negative values for clean firms are likely driven by the significant stock price drops for distressed firms. When we look at the three alternative measures of insider trading activities (*CHNSV2*, *CHNSP*, and *CHNSF* as discussed in Section V), which are less likely to be mechanically affected by stock price, their mean is –0.049, 0.028, and 1.276, respectively, for clean firms.[13]

In addition, fewer firms in the GC group are audited by Big N auditors, while GC firms have a higher frequency of losses, a lower Altman's Z-score, a longer period of announcement lag, and a lower frequency of external financing in the next year. Furthermore, GC firms have lower stock

---

[10] All results are both quantitatively and qualitatively similar when variables are not winsorized.

[11] The reason that the mean of the predicted changes differs from that of the actual changes is that a much larger sample is used in the first-stage prediction model as noted in Appendix A. The negative mean of *CHNSV* is mainly driven by the significant stock price drops for distressed firms over time. When *CHNSV* is measured in shares rather than in dollar amount, the mean of *CHSNV* is statistically indistinguishable from zero. In an untabulated test, we replace *PREDICT_CHNSV* with *CHNSV* in Equation (1) and find similar results.

[12] Given that Ma (2001) and Nasser and Gup (2008) find a lower level of insider sales before bankruptcy announcements, these results may not be surprising.

[13] For the matched sample as discussed in Section V, the mean of *PREDICT_CHNSV* and *CHNSV* is 0.197 and –0.051 for clean firms, respectively. The means of the three alternative insider-trading measures (*CHNSV2*, *CHNSP*, and *CHNSF*) are 0.008, 0.052, and 0.777, respectively, for the matched clean firms.

TABLE 1

Descriptive Statistics

Panel A: Descriptive Statistics for the Full Sample

Variable	n	Mean	Median	Std. Dev.	Q1	Q3
PREDICT_CHNSV	12,329	0.167	0.135	1.159	−0.552	0.868
CHNSV ($000)	12,329	−133.56	0	202.47	−196.35	85.49
CHNSV (in log, $000)	12,329	−0.214	0	4.115	−2.512	1.889
BIGN	12,329	0.736	1	0.440	0	1
LOSS	12,329	0.869	1	0.336	1	1
ZSCORE	12,329	0.575	1.034	5.150	−0.147	2.238
ANNLAG	12,329	60.931	55	36.581	38	80
NEWFINANCE	12,329	0.581	1	0.494	0	1
RETURN	12,329	0.023	−0.163	0.966	−0.501	0.241
VOLATILITY	12,329	0.207	0.126	0.175	0.120	0.258
SIZE	12,329	5.038	4.801	1.970	3.676	6.173
LEV	12,329	0.512	0.471	0.418	0.247	0.712
CLEV	12,329	0.052	0.026	0.297	−0.011	0.094
AGE	12,329	2.335	2.197	0.715	1.791	2.833
OCF	12,329	−0.108	−0.021	0.309	−0.141	0.038
INVESTMENT	12,329	0.278	0.174	0.276	0.041	0.461
Clients' Economic Significance Measures						
DEPENDENCE (by audit fee ratio)	11,240	0.121	0.049	0.203	0.022	0.114
DEPENDENCE (by market value ratio)	12,329	0.115	0.041	0.287	0.018	0.102
Auditors' Litigation Concern Measure						
FRAUD	12,329	0.214	0	0.410	0	0
Auditors' Reputation Measures						
EXPERT	12,030	0.467	0	0.499	0	1
OFFICESIZE	11,240	9.156	9.412	1.903	8.019	10.539
Audit Committee Independence Measure						
INDEP	2,060	0.763	1	0.420	1	1

Panel B: Descriptive Statistics of Firm Characteristics by Audit Opinion

	Clean Firms (n = 11,528)		Going-Concern Firms (n = 801)		p-values of the Difference	
Variable	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
PREDICT_CHNSV	0.189	0.152	−0.131	−0.142	< 0.01	< 0.01
CHNSV ($000)	−114.25	0	−319.59	0	< 0.01	< 0.01
CHNSV (in log, $000)	−0.191	0	−0.505	0	< 0.01	< 0.01
BIGN	0.741	1	0.659	1	< 0.01	< 0.01
LOSS	0.861	1	0.981	1	< 0.01	< 0.01
ZSCORE	0.859	1.132	−3.524	−1.164	< 0.01	< 0.01
ANNLAG	59.107	53	87.186	90	< 0.01	< 0.01
NEWFINANCE	0.595	1	0.488	1	< 0.01	< 0.01
RETURN	0.058	−0.125	−0.443	−0.603	< 0.01	< 0.01
VOLATILITY	0.198	0.166	0.305	0.256	< 0.01	< 0.01

(continued on next page)

TABLE 1 (continued)

Variable	Clean Firms (n = 11,528) Mean	Clean Firms (n = 11,528) Median	Going-Concern Firms (n = 801) Mean	Going-Concern Firms (n = 801) Median	p-values of the Difference Mean	p-values of the Difference Median
SIZE	5.113	4.862	3.958	3.693	< 0.01	< 0.01
LEV	0.493	0.456	0.795	0.698	< 0.01	< 0.01
CLEV	0.039	0.023	0.244	0.122	< 0.01	< 0.01
AGE	2.345	2.197	2.183	2.079	< 0.01	< 0.01
OCF	−0.087	−0.021	−0.409	−0.162	< 0.01	< 0.01
INVESTMENT	0.284	0.184	0.196	0.090	< 0.01	< 0.01

(continued on next page)

returns, higher return volatility, smaller size, higher leverage and change in leverage, lower operating cash flow, and a lower level of investment in securities as a percentage of total assets. All these differences are as expected and highlight the importance of controlling for these characteristics when estimating Equation (1). To address the concerns that the inclusion of these variables cannot fully control for the differences between the two groups, we report the results based on a matched sample in Section V.

Table 1, Panel C reports Pearson and Spearman correlations below and above the diagonal, respectively. The correlations between the control variables and the predicted change in insider sales (*PREDICT_CHNSV*) are generally quite low, with a magnitude of 0.131 or less.

Multivariate Analysis

Table 2 reports the results from testing H1 based on a two-step regression method. We find that the predicted change in insider selling activity (*PREDICT_CHNSV*) is negatively associated with the likelihood of receiving a going-concern opinion (coefficient = −0.136, $p < 0.01$). The evidence suggests that auditors are less likely to issue going-concern opinions for firms when insiders sell abnormally more shares. A one standard deviation increase in the predicted change of net insider sales, which translates into an increase of roughly $60,250 in the change of net insider sales, reduces the likelihood of receiving a going-concern opinion by 1.39 percent, which is around 21 percent of the mean going-concern rate for our sample. The economic magnitude of 1.39 percent is comparable to other factors documented in prior studies, such as 1.26 percent for auditor office size (Francis and Yu 2009) and 1 percent for total fee ratios (Li 2009). The marginal effect of the change in the level of insider selling compares with that of other determinants of going-concern opinions, such as firm size (−1.22 percent), stock returns (−1.35 percent), leverage (1.25 percent), and Big N auditors (1.24 percent).[14]

The coefficients on the control variables are generally in the predicted directions. Big N auditors are more likely to issue going-concern opinions (Mutchler et al. 1997). On the one hand, auditors are more likely to issue going-concern opinions for firms with losses, longer announcement lags, higher return volatility, and higher leverage. On the other hand, firms are less likely to receive

[14] The model performs well in predicting going-concern opinions as shown by the high value of concordance (0.904). When the model is estimated without the measure of insider selling activity (i.e., *PREDICT_CHNSV*), the concordance value reduces to 0.900. Though *PREDICT_CHNSV* brings an incremental concordance of only 0.004, it is comparable to that of other control variables in Equation (1), such as *SIZE*, *LOSS*, *AGE*, *LEV*, *OCF*, and *BIGN*, whose incremental concordance ranges from 0.001 to 0.006.

TABLE 1 (continued)

Panel C: Pearson and Spearman Correlation Coefficients (n = 12,329)

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
(1) *PREDICT_CHNSV*		−0.014	**−0.128**	**0.084**	0.003	0.019	**−0.027**	**−0.138**	0.032	**−0.051**	**−0.067**	0.017	0.002	0.013	**−0.124**
(2) *BIGN*	−0.021		**0.039**	**−0.019**	**−0.291**	**0.088**	**0.037**	−0.013	**0.299**	−0.002	0.000	**−0.047**	**0.031**	**0.126**	**−0.046**
(3) *LOSS*	**−0.131**	**0.039**		**−0.186**	**0.038**	−0.013	**−0.156**	**0.214**	**−0.119**	**−0.084**	**0.195**	**−0.123**	**0.078**	**0.164**	**0.087**
(4) *ZSCORE*	**0.097**	−0.016	**−0.113**		**−0.061**	**0.027**	**0.125**	**−0.202**	**0.147**	**−0.359**	**−0.323**	**0.132**	**0.415**	**−0.173**	**−0.227**
(5) *ANNLAG*	0.015	**−0.204**	**0.034**	**−0.042**		**−0.172**	**−0.065**	**0.029**	**−0.382**	**0.084**	**0.083**	**0.095**	**−0.077**	**−0.188**	**0.218**
(6) *NEWFINANCE*	0.022	**0.102**	−0.012	0.001	**−0.144**		**0.144**	**0.058**	**0.061**	**−0.065**	**−0.059**	**−0.068**	−0.002	**0.041**	**−0.105**
(7) *RETURN*	**−0.025**	**0.037**	**−0.059**	**0.034**	**−0.048**	**0.115**		**−0.184**	**0.067**	**−0.037**	**−0.162**	**0.131**	**0.141**	0.003	**−0.178**
(8) *VOLATILITY*	**−0.125**	**−0.016**	**0.163**	**−0.159**	**0.019**	**0.038**	**0.089**		**−0.295**	**−0.097**	**0.111**	**−0.300**	**−0.284**	**0.185**	**0.186**
(9) *SIZE*	0.023	**0.286**	**−0.151**	**0.178**	**−0.255**	**0.069**	**−0.033**	**−0.258**		**0.326**	**−0.040**	**0.129**	**0.407**	**−0.281**	**−0.148**
(10) *LEV*	**−0.056**	−0.004	**−0.038**	**−0.284**	**0.099**	**−0.053**	**−0.032**	**−0.024**	**0.213**		**0.346**	**0.183**	**0.191**	**−0.496**	**0.149**
(11) *CLEV*	**−0.051**	−0.004	**0.094**	**−0.234**	**0.068**	**−0.035**	**−0.061**	**0.053**	**−0.052**	**0.677**		**−0.073**	**−0.098**	**−0.064**	**0.182**
(12) *AGE*	0.017	**−0.034**	**−0.120**	**0.068**	**0.052**	**−0.062**	**0.032**	**−0.256**	**0.191**	**0.117**	**−0.039**		**0.217**	**−0.246**	**−0.053**
(13) *OCF*	0.003	**0.025**	**−0.057**	**0.555**	**−0.060**	0.006	**0.068**	**−0.226**	**0.372**	**−0.027**	**−0.174**	**0.190**		**−0.426**	**−0.147**
(14) *INVESTMENT*	0.018	**0.121**	**0.158**	**−0.103**	**−0.137**	**0.049**	**0.047**	**0.129**	**−0.287**	**−0.316**	**−0.037**	**−0.251**	**−0.360**		**−0.083**
(15) *GC OPINION*	**−0.126**	**−0.046**	**0.087**	**−0.210**	**0.189**	**−0.126**	**−0.131**	**0.185**	**−0.145**	**0.178**	**0.170**	**−0.056**	**−0.257**	**−0.078**	

Table 1 reports summary statistics for the full sample in Panel A and subsamples partitioned by the two types of audit opinions in Panel B. Panel C presents Pearson (below the diagonal) and Spearman (above the diagonal) correlations, with the correlation coefficients with a significance level of 0.05 or better in bold. All continuous variables, except stock returns (*RETURN*), are winsorized at the 1st and 99th percentile.

All variables are defined in Appendix C.

TABLE 2

Insider Trading Incentives and the Likelihood of a Going-Concern Opinion

$$\begin{aligned} Pr(GC\ OPINION_t = 1) = {} & \beta_0 + \beta_1(PREDICT\_CHNSV_t) + \beta_2(ZSCORE_t) + \beta_3(LOSS_t) \\ & + \beta_4(SIZE_t) + \beta_5(AGE_t) + \beta_6(RETURN_t) + \beta_7(VOLATILITY_t) \\ & + \beta_8(LEV_t) + \beta_9(CLEV_t) + \beta_{10}(OCF_t) + \beta_{11}(ANNLAG_t) \\ & + \beta_{12}(INVESTMENTS_t) + \beta_{13}(NEWFINANCE_t) + \beta_{14}(BIGN_t) \\ & + \sum Industries + \sum Years + \varepsilon_t. \end{aligned} \quad (1)$$

Variables	**Coeff. Estimate**	**Wald Chi-Square**	**Marginal Effects**
Intercept	−3.608***	41.24	
PREDICT_CHNSV	−0.136***	12.99	−1.39%
ZSCORE	−0.093***	31.15	−3.13%
LOSS	1.582***	26.98	2.27%
SIZE	−0.177***	31.75	−1.22%
AGE	−0.101*	3.04	−1.05%
RETURN	−0.848***	48.28	−1.35%
VOLATILITY	3.011***	103.53	1.12%
LEV	0.845***	22.52	1.25%
CLEV	0.446	2.52	0.54%
OCF	−1.264***	32.91	−3.25%
ANNLAG	0.016***	157.97	3.49%
INVESTMENT	−2.443***	76.86	−4.78%
NEWFINANCE	−0.457***	26.44	−1.98%
BIGN	0.226**	4.87	1.24%
Year fixed effects		Yes	
Industry fixed effects		Yes	
Pseudo R^2		0.348	
Percent Concordant		0.904	
n		12,329	

*, **, *** Denote significance at the 0.10, 0.05, and 0.01 levels, respectively.
This table reports the results of the second-stage regression testing the association between insider trading and the likelihood of receiving a going-concern opinion. *PREDICT_CHNSV* is the predicted change in insider sales obtained from the first-stage regression Equation (A1) as shown in Appendix A. Robust standard errors clustered at the client firm and auditor levels are used to derive the Wald statistics.
All variables for Equation (1) are defined in Appendix C.

going-concern opinions when they are larger and older and when they have higher Z-scores, new financing activities, higher past stock returns, higher operating cash flows, and larger amounts of cash and investment securities. In summary, the results in Table 2 support the prediction in H1 that abnormally high insider sales are associated with a lower likelihood of receiving a going-concern opinion.

Cross-Sectional Analysis

To test H2 through H5, we conduct four cross-sectional analyses. Specifically, we augment Equation (1) by including an additional variable that measures clients' economic significance,

auditors' litigation concerns, auditors' reputation, and audit committee independence, alternately, and their interaction with *PREDICT_CHNSV*. Table 3 reports the results of these analyses.

Economic Significance of Clients

Panel A of Table 3 presents the results from testing H2 by adding a variable for auditor economic dependence (*DEPENDENCE*). Following prior studies (e.g., Reynolds and Francis 2000), we proxy auditor economic dependence by fee dependence at the auditor local office level with descriptive statistics presented in Panel A of Table 1. Specifically, we gauge the economic significance of clients as the ratio of a specific client's audit fee (or market value) to the total audit fees (or total market value) for all the clients of an incumbent auditor's local office, where fees and clients' market values are in natural logarithm. As shown in Panel A of Table 3, the coefficient on *PREDICT_CHNSV* continues to load negatively at $p < 0.05$. Our focus, the interaction term *DEPENDENCE* * *PREDICT_CHNSV*, loads negatively and is statistically significant at $p < 0.05$ for both measures of clients' economic significance. Thus, we find evidence consistent with H2.

Auditors' Litigation Concerns

Panel B of Table 3 presents the results from testing H3 by adding a variable for auditor litigation risk (*FRAUD*). Previous research provides evidence that fraud conducted by clients increases auditors' litigation risk (e.g., Palmrose 1988; Carcello and Palmrose 1994; Bonner et al. 1998). Auditors likely face a higher risk of litigation when their clients announce any type of fraud.[15] Specifically, we define the indicator variable, *FRAUD*, as 1 if any of a local audit office's clients announced a financial fraud in the most recent two years, and 0 otherwise. Panel A of Table 1 shows that, for 21.4 percent of firm years, the incumbent auditors face a higher risk of litigation in relation to financial fraud announced by clients. The main effect of *PREDICT_CHNSV* continues to be negative and significant. Consistent with H3, the interaction of *PREDICT_CHNSV* with *FRAUD* has a positive effect on the issuance of going-concern opinions at the 0.05 significance level. These results suggest that auditors are more independent and conservative in issuing going-concern reports when facing a higher risk of litigation. A Wald test ($\beta_1 + \beta_3 = 0$) rejects the null (Chi-square $= 4.78$, $p = 0.03$), suggesting that insider trading incentives still have an effect on the issuance of going-concern opinions even when auditors have a high risk of litigation, although the effect is much smaller.

Auditors' Reputation Concerns

Panel C of Table 3 presents the results from testing H4 by adding a variable for auditor reputation (*EXPERT* and *OFFICESIZE*). Following prior studies, we identify an auditor as an expert in a certain industry if the auditor has a dominant market share in that industry. Specifically, *EXPERT* is coded 1 if the incumbent auditor ranks as a top-two firm in the client's industry (two-digit SIC) in terms of market share of sales, and 0 otherwise. Results are similar when we define *EXPERT* as 1 for the top three industry leaders in the audit market.[16] Aside from industry specialization, we use office size as an alternative measure of auditor reputation. Francis and Yu (2009) show that the office size of Big 4 auditors is positively associated with audit quality. Following their study, we measure local office size (*OFFICESIZE*) by the natural logarithm of total audit fees. Panel A of Table 1 presents descriptive statistics for *EXPERT* and *OFFICESIZE*.

---

[15] Accounting frauds are irregularities as defined in Hennes et al. (2008). We thank Andrew Leone for sharing the dataset of all GAO restatements at his personal website: http://sbaleone.bus.miami.edu/.

[16] Prior studies provide evidence that auditor reputation in regard to industry specialization is related to higher audit quality (e.g., Craswell et al. 1995; Dunn and Mayhew 2002; Chung and Kallapur 2003; Lim and Tan 2008).

TABLE 3

Cross-Sectional Analysis of the Association between Insider Trading and the Likelihood of a Going-Concern Opinion

$$\begin{aligned} Pr(GC\ OPINION_t = 1) = {} & \beta_0 + \beta_1(PREDICT\_CHNSV_t) + \beta_2(EXPERIMENT_t) \\ & + \beta_3(EXPERIMENT_t{*}PREDICT\_CHNSV_t) + \beta_4(ZSCORE_t) \\ & + \beta_5(LOSS_t) + \beta_6(SIZE_t) + \beta_7(AGE_t) + \beta_8(RETURN_t) \\ & + \beta_9(VOLATILITY_t) + \beta_{10}(LEV_t) + \beta_{11}(CLEV_t) + \beta_{12}(OCF_t) \\ & + \beta_{13}(REPORTLAG_t) + \beta_{14}(INVESTMENTS_t) \\ & + \beta_{15}(NEWFINANCE_t) + \beta_{16}(BIGN_t) + \sum Industries \\ & + \sum Years + \varepsilon_t. \end{aligned} \tag{1}$$

Panel A: Clients' Economic Significance

	Clients' Economic Significance Measure					
	Audit Fee Ratio			Market Value Ratio		
Variables	**Coeff. Estimate**	**Wald Chi-Square**	**Marginal Effects**	**Coeff. Estimate**	**Wald Chi-Square**	**Marginal Effects**
Intercept	−3.044***	11.42		−2.988***	23.48	
PREDICT_CHNSV	−0.082**	5.96	−0.87%	−0.076**	6.05	−0.79%
DEPENDENCE	−0.288	1.35	−1.15%	−0.312**	4.06	−1.22%
*DEPENDENCE * PREDICT_CHNSV*	−0.205**	3.92	−2.45%	−0.226**	4.22	−2.51%
Control variables		Yes			Yes	
Year fixed effects		Yes			Yes	
Industry fixed effects		Yes			Yes	
Pseudo R^2		0.349			0.351	
Percent Concordant		0.902			0.907	
n		11,240			12,329	

Panel B: Auditor Litigation Concern

Variables	**Coeff. Estimate**	**Wald Chi-Square**	**Marginal Effects**
Intercept	−3.078***	33.28	
PREDICT_CHNSV	−0.195***	7.54	−1.67%
FRAUD	0.088*	3.56	0.78%
*FRAUD * PREDICT_CHNSV*	0.103**	4.27	0.88%
Control variables		Yes	
Year fixed effects		Yes	
Industry fixed effects		Yes	
Pseudo R^2		0.351	
Percent Concordant		0.906	
n		12,329	

(continued on next page)

TABLE 3 (continued)

Panel C: Auditor Reputation Concern

	Auditor Reputation Measures					
	EXPERT			*OFFICESIZE*		
Variables	**Coeff. Estimate**	**Wald Chi-Square**	**Marginal Effects**	**Coeff. Estimate**	**Wald Chi-Square**	**Marginal Effects**
Intercept	−3.155***	12.77		−2.867***	21.36	
PREDICT_CHNSV	−0.188**	5.87	−1.62%	−0.172**	6.33	−1.58%
EXPERT	0.036	0.98	0.25%			
EXPERT ∗ *PREDICT_CHNSV*	0.106**	4.21	1.18%			
OFFICESIZE				0.065**	4.06	0.48%
OFFICESIZE ∗ *PREDICT_CHNSV*				0.027**	4.35	0.96%
Control variables		Yes			Yes	
Year fixed effects		Yes			Yes	
Industry fixed effects		Yes			Yes	
Pseudo R^2		0.351			0.348	
Percent Concordant		0.906			0.904	
n		12,030			11,240	

Panel D: Audit Committee Independence

Variables	**Coeff. Estimate**	**Wald Chi-Square**	**Marginal Effects**
Intercept	−11.979***	25.05	
PREDICT_CHNSV	−0.302**	4.07	−1.70%
INDEP	0.473	1.19	0.42%
INDEP ∗ *PREDICT_CHNSV*	0.279**	4.04	1.63%
Control variables		Yes	
Year fixed effects		Yes	
Industry fixed effects		Yes	
Pseudo R^2		0.235	
Percent Concordant		0.844	
n		2,060	

*, **, *** Denote significance at the 0.10, 0.05, and 0.01 levels, respectively.
This table reports the results of the second-stage regression testing the effect of clients' economic significance, auditor litigation concern and reputation, and audit committee independence on the association between insider trading and the likelihood of receiving a going-concern opinion. The variable, *EXPERIMENT*, in the regression equation takes alternative measures of clients' economic significance (*DEPENDENCE*), auditor litigation risk (*FRAUD*), auditor reputation (*EXPERT*), and audit committee independence (*INDEP*). Robust standard errors clustered at the client firm and auditor levels are used to derive the Wald statistics.
All variables, including the two measures of *DEPENDENCE* (audit fee ratio and market value ratio), are defined in Appendix C.

Consistent with our expectation, the results in Panel C of Table 3 show a positive coefficient on the interaction between *PREDICT_CHNSV* and *EXPERT*, significant at $p < 0.05$. A test of ($\beta_1 + \beta_3 = 0$) rejects the null with a p-value of 0.08, suggesting that the effect of insider trading on going-concern opinions still exists for firms with reputable auditors, notwithstanding a significantly

smaller effect. Similarly, the interaction between *PREDICT_CHNSV* and *OFFICESIZE* also loads positively and is statistically significant at the 0.05 level.[17] The test of ($\beta_1 + \beta_3 = 0$) is rejected with a p-value of 0.02. To ease interpretation, we create an indicator *BIGOFFICE* for those firms whose auditors' local office size is larger than the sample median and replace *OFFICESIZE* with *BIGOFFICE* in the regression. We find a positive coefficient on the interaction *BIGOFFICE* * *PREDICT_CHNSV*. A test of ($\beta_1 + \beta_3 = 0$) rejects the null with a p-value of 0.07, suggesting that insiders' selling incentives still affect the issuance of going-concern opinions by reputable auditors. Taken together, the results imply that insider trading incentives have a smaller effect on the likelihood of receiving a going-concern opinion when auditors are more reputable.

Audit Committee Independence

Panel D of Table 3 reports the results from testing H5 by adding a variable for audit committee independence (*INDEP*). We obtain audit committee information from the Investor Responsibility Research Center (IRRC). Note the definition of independent director by the IRRC is more restrictive than the one set forth by the SEC rule and stock exchanges such as NYSE and NASDAQ.[18] It is unclear *ex ante* whether such a discrepancy will compromise board quality and provide managers with opportunities to influence auditors. We thus follow prior studies and abide by the IRRC definition of independent director (Duchin et al. 2010; Guo and Masulis 2012). An indicator variable, *INDEP*, is defined as 1 if a client's audit committee is fully represented by independent directors, and 0 otherwise. As reported in Panel A of Table 1, 76.3 percent of the 2,060 firm years have a fully independent audit committee.[19]

Table 3, Panel D shows a positive but insignificant coefficient on *INDEP*, suggesting that audit committee independence *per se* does not affect the likelihood of auditors issuing going-concern opinions. Consistent with our prediction in H5, the coefficient on the interaction *INDEP* * *PREDICT_CHNSV* is positive and statistically significant at $p < 0.05$. A further test of ($\beta_1 + \beta_3 = 0$) fails to reject the null with a p-value of 0.323, suggesting that the insider trading effect on going-concern opinion issuance is muted for firms with fully independent audit committees.

V. ADDITIONAL ANALYSIS, SENSITIVITY TESTS, AND FURTHER DISCUSSION

Endogeneity Issues

A System of Two Structural Models Estimation

Our analyses of insider trading use the predicted change in insider sales to control for endogeneity between insider trading and auditor opinion. In addition, the cross-sectional analyses in

---

[17] We also use *BIGN* as an alternative proxy for auditor reputation, where *BIGN* is as previously defined. The coefficient on the interaction of *BIGN* with *PREDICT_CHNSV* is positive and statistically insignificant at the 0.10 level.

[18] For example, according to the IRRC, a director who is a former employee of the firm is not independent, even if the employment terminated more than three years before the director was seated. Another example is that a director who has business relations with the firm is not independent according to the IRRC, even if the business relationship is insignificant. In contrast, Rule 10A-3 (SOX), NYSE, and NASDAQ allow former employees to become independent directors if more than three years have passed since their employment and also allow independent directors to have business relations with the firm as long as the transactions are not significant.

[19] Alternatively, we use audit committee members' trading activities in their own firms' stocks to gauge audit committee independence. We calculate the net trading activities as the number of sale transactions minus the number of purchase transactions for all audit committee members over a fiscal year. We then define a dummy variable equal to 1 if the net trading activities are positive, and 0 otherwise. Based on this sample, we find the audit committee members' selling activity strengthens the negative relation between insider trading and going-concern opinions (untabulated).

the previous section provide further support that insiders' trading incentives influence the likelihood of auditors issuing going-concern reports. Auditors with greater litigation concerns and higher reputation are more likely to issue going-concern opinions, all else being equal. If only the reduction of insider selling in anticipation of a forthcoming going-concern opinion drives our results, then we would expect the results to be more pronounced for auditors with greater litigation concerns and higher reputation and for clients with fully independent audit committees. The evidence, however, points in the opposite direction, which is consistent with the likelihood that managers discourage auditors from issuing going-concern opinions when they sell abnormally more shares.

To further illuminate this issue, we estimate a system of two structural models. One equation models auditor opinion based on Equation (1), using *CHNSV* instead of *PREDICT_CHNSV* as the endogenous independent variable. The other equation models insider trading based on Equation (A1) in Appendix A with three additional variables: contemporaneous auditor opinion (GC_t), which is the endogenous independent variable, current year's accumulated stock returns (RET_t), and the change in return-on-equity for the current year (ΔROE_t). Following DeFond et al. (2002), we use a probit regression to estimate the auditor opinion model and an OLS regression to estimate the insider-trading model. The untabulated results show that the coefficient on *GC* is positive but not statistically significant ($p = 0.57$) in the insider trading model. In contrast, the coefficient on *CHNSV* is significantly negative ($p < 0.01$) in the auditor opinion model. Therefore, our results are robust to the use of a system of two equations.

Auditor Switch Analysis

A natural extension is to investigate whether auditors benefit from acquiescing to management when insiders have abnormally high stock sales. Using a sample of 10,494 clean firm years, we estimate a logistic model by regressing an indicator variable, *SWITCH*, on *PREDICT_CHNSV* and a set of control variables.[20] The indicator *SWITCH* is equal to 1 if an auditor change occurs in the next fiscal year, and 0 otherwise. Results (untabulated) show that among the distressed firms receiving clean opinions, 12.8 percent experience an auditor change during the next fiscal year. More importantly, untabulated results show that the variable of interest, *PREDICT_CHNSV*, has a significantly negative coefficient ($p < 0.01$) suggesting that clean firms with higher levels of abnormal insider sales are less likely to have an auditor switch in the future. The evidence is consistent with the notion that auditors can retain clients by acquiescing to management pressure but inconsistent with the alternative explanation that insiders reduce their selling in anticipation of going-concern reports.

Pre- and Post-SOX

Auditor independence increased in various ways after the 2002 Sarbanes-Oxley Act.[21] Recent research documents that auditors, particularly in non-Big N firms, became more conservative in issuing going-concern reports after SOX and that audit committees became more independent

[20] The set of control variables include change in firm size and change in operating cash flows from $t-1$ to t, external financing, change in times-interest-earned from $t-1$ to t, industry market share of an auditor in the client industry, Altman's Z-score, leverage, loss indicator, industry fixed effects and year fixed effects.

[21] First, the act prohibited auditors from providing nonaudit consulting services, with the aim of enhancing auditors' independence. Second, SOX empowered federal courts and the SEC to impose equitable remedies for violations of federal securities laws, which suggests civil monetary penalties against auditors increased following SOX (Rashkover and Winter 2005). Third, the creation of the PCAOB could increase auditors' exposure to legal liability *vis-à-vis* penalties for audit-related violations. Fourth, SOX requires internal audit committees to take more responsibility to ensure the autonomy of external auditors without being influenced by management.

(Hoitash and Hoitash 2008; Myers et al. 2011). In addition, prior studies find that insiders' purchases and sales dropped considerably and that insiders became more careful in exploiting negative information after SOX (Chang et al. 2011; Lee et al. 2011). If auditor independence increased while insiders trading on private information decreased after SOX, then we would expect the negative relation as stated in H1 to become weaker after SOX.

Different sections of SOX took effect on various dates. For example, Section 301 on auditor independence became effective on May 6, 2003, while the compliance for Section 404 and related disclosures were required for fiscal-years ending after November 15, 2004. Following prior studies (Iliev 2010; Li et al. 2010; Krishnan et al. 2011), we partition our sample period into pre-SOX period (2000 through 2003) and post-SOX period (2004 through 2007).[22] We then estimate Equation (1) for the two sub-periods jointly in a single regression, allowing all coefficients to differ across the two periods and test the difference in the coefficient estimates across the two periods. Results reported in Table 4 show that the estimated coefficient on *PREDICT_CHNSV* is negative and statistically significant both before and after SOX. However, the absolute magnitude is much smaller after SOX, and the difference in the coefficient estimate on *PREDICT_CHNSV* is statistically significant ($p < 0.05$). In addition, the coefficient estimate on *BIGN* increases significantly after SOX, suggesting that Big N auditors are more likely to issue going-concern opinions after SOX. Taken as a whole, these results indicate that auditor independence increased after SOX.

Sensitivity Tests

Alternative Sample

Clean-opinion firms differ from going-concern firms, as demonstrated in Panel B of Table 1. Thus, although we include a comprehensive set of control variables in the analysis, one may be still concerned that our results are driven by correlated omitted variables. We address this concern using a matching procedure to find a sample of clean firms that have similar fundamental characteristics as the GC firms in year *t*. More specifically, we match a clean firm with a GC firm in the current year by first sorting all firms into deciles based on annual stock return and Altman's Z-score independently. We then select all clean firms that are in the same decile of stock return and in the same decile of Altman's Z-score as the GC firm. Among these clean firms, we calculate their matching scores that are the sum of their distances to the GC firm in terms of stock performance and Altman's Z-score, respectively, and select the one with the smallest matching score.[23] This matching procedure yields 801 unique clean firms matched for the 801 *GC* firms.

The comparison of the fundamentals between the two groups is tabulated in Panel A of Table 5. We find no significant difference between the matched clean firms and GC firms in stock returns, Altman's Z-score and firm size, which are three major determinants of auditors' issuance of GC opinions and insiders' decision to trade. There is no significant difference in other firm characteristics between GC firms and the matched clean firms such as return volatility, operating cash flows, and the indicators for Big N auditors and for loss firms. Overall, our matching procedure seems to work pretty well in identifying clean firms with similar fundamentals as the GC firms.

---

[22] Our results are both quantitatively and qualitatively similar when we use 2003 as the cutoff for the pre- and the post-SOX period.

[23] The distance of stock performance (Altman's Z-score) is measured as the absolute value of the difference in annual stock returns (Altman's Z-scores) between a clean firm and the GC firm, scaled by the absolute value of the GC firm's annual stock return (Altman's Z-score).

TABLE 4

Insider Trading and the Likelihood of Receiving a Going-Concern Opinion before and after SOX

$$\begin{aligned} Pr(GC\ OPINION_t = 1) = {} & \beta_0 + \beta_1(PREDICT\_CHNSV_t) + \beta_2(ZSCORE_t) + \beta_3(LOSS_t) \\ & + \beta_4(SIZE_t) + \beta_5(AGE_t) + \beta_6(RETURN_t) + \beta_7(VOLATILITY_t) \\ & + \beta_8(LEV_t) + \beta_9(CLEV_t) + \beta_{10}(OCF_t) + \beta_{11}(ANNLAG_t) \\ & + \beta_{12}(INVESTMENTS_t) + \beta_{13}(NEWFINANCE_t) + \beta_{14}(BIGN_t) \\ & + \sum Industries + \sum Years + \varepsilon_t. \end{aligned} \tag{1}$$

	Pre-SOX		Post-SOX		Pre- versus Post-SOX	
Variables	**Coeff. Estimate**	**Wald Chi-Square**	**Coeff. Estimate**	**Wald Chi-Square**	**Coeff. Difference**	**Wald Chi-Square**
Intercept	−3.285***	22.42	−2.955***	18.45	0.330	1.15
PREDICT_CHNSV	−0.171***	8.55	−0.084*	2.94	0.087**	4.12
ZSCORE	−0.112***	29.94	−0.055*	3.89	0.057*	3.03
LOSS	1.452***	16.52	1.968***	9.24	0.516	0.46
SIZE	−0.066***	8.11	−0.254***	10.37	−0.189**	4.02
AGE	−0.137*	3.09	−0.136	0.96	0.001	0.00
RETURN	−0.676***	31.67	−1.154***	12.45	−0.478*	2.81
VOLATILITY	2.642***	70.89	6.074***	39.65	3.432***	7.28
LEV	0.635***	10.45	1.203***	11.27	0.568*	2.98
CLEV	0.381	1.37	0.186	0.25	−0.195	0.21
OCF	−1.317***	29.51	−1.888**	6.98	−0.571	1.87
ANNLAG	0.021***	98.23	0.013***	38.86	−0.008***	10.72
INVESTMENT	−2.884***	69.42	−1.893***	7.35	−0.991**	3.94
NEWFINANCE	−0.656***	61.04	−0.675***	11.32	−0.019	0.85
BIGN	0.181*	3.35	0.613***	7.28	0.432**	4.58
Year fixed effects			Yes			
Industry fixed effects			Yes			
Pseudo R^2			0.352			
Percent Concordant			0.908			
n			12,329			

*, **, *** Denote significance at the 0.10, 0.05, and 0.01 levels, respectively.
This table reports the results of the second-stage regression based on Equation (1) for the pre-SOX (2000 through 2003) and post-SOX (2004 through 2007) periods and the differences across these two periods. Robust standard errors clustered at the client firm and auditor levels are used to derive the Wald statistics.
All variables are defined in Appendix C.

Based on this matched sample, we re-run the main analysis; the results are reported in Panel B of Table 5. As can be seen, the coefficient on our main variable of interest, *PREDICT_CHNSV*, continues to be significantly negative ($p < 0.05$). Moreover, the marginal effect of the change in net insider sales is much higher compared to that based on the full sample as reported in Table 2. Overall, our results are less likely to be driven by correlated omitted firm characteristics.

TABLE 5

Insider Trading and the Likelihood of Receiving a Going-Concern Opinion based on a Matched Sample

$$\begin{aligned} Pr(GC\ OPINION_t = 1) = {} & \beta_0 + \beta_1(PREDICT\_CHNSV_t) + \beta_2(ZSCORE_t) + \beta_3(LOSS_t) \\ & + \beta_4(SIZE_t) + \beta_5(AGE_t) + \beta_6(RETURN_t) + \beta_7(VOLATILITY_t) \\ & + \beta_8(LEV_t) + \beta_9(CLEV_t) + \beta_{10}(OCF_t) + \beta_{11}(ANNLAG_t) \\ & + \beta_{12}(INVESTMENTS_t) + \beta_{13}(NEWFINANCE_t) + \beta_{14}(BIGN_t) \\ & + \sum Industries + \sum Years + \varepsilon_t. \end{aligned} \tag{1}$$

Panel A: Comparison in Firm Characteristics between GC Firms and Matched Clean Firms

	Matched Clean Firms (n = 801)		Going-Concern Firms (n = 801)		p-values of the Difference	
Variable	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
BIGN	0.678	1	0.659	1	0.23	0.19
LOSS	0.979	1	0.981	1	0.57	0.37
ZSCORE	−3.631	−1.227	−3.524	−1.164	0.18	0.21
ANNLAG	60.298	58	87.186	90	< 0.01	< 0.01
NEWFINANCE	0.755	1	0.488	1	< 0.01	< 0.01
RETURN	−0.432	−0.627	−0.443	−0.603	0.67	0.60
VOLATILITY	0.285	0.259	0.305	0.256	0.32	0.26
SIZE	4.022	3.721	3.958	3.693	0.35	0.31
LEV	0.630	0.534	0.795	0.698	< 0.01	< 0.01
CLEV	0.171	0.091	0.244	0.122	< 0.01	< 0.01
AGE	2.035	1.946	2.183	2.079	< 0.01	< 0.01
OCF	−0.500	−0.278	−0.409	−0.162	0.34	0.12
INVESTMENT	0.379	0.346	0.196	0.090	< 0.01	< 0.01

Panel B: Regression Results

	PREDICT_CHNSV		
Variables	**Coeff. Estimate**	**Wald Chi-Square**	**Marginal Effects**
Intercept	0.789	1.06	
PREDICT_CHNSV	−0.155**	6.09	−14.92%
ZSCORE	−0.009	0.17	−0.35%
LOSS	1.213*	3.44	13.41%
SIZE	−0.104	2.29	−4.26%
AGE	−0.279**	4.61	−9.27%
RETURN	−0.172	1.14	−7.03%
VOLATILITY	1.669*	3.32	2.64%
LEV	0.713***	8.29	13.77%
CLEV	−0.027	0.01	−0.24%
OCF	−0.686*	3.73	−9.71%
ANNLAG	0.017***	39.39	33.49%

(continued on next page)

TABLE 5 (continued)

	PREDICT_CHNSV		
Variables	**Coeff. Estimate**	**Wald Chi-Square**	**Marginal Effects**
INVESTMENT	−2.989***	54.36	−30.27%
NEWFINANCE	−1.241***	60.64	−24.06%
BIGN	0.423***	7.98	13.03%
Year fixed effects		Yes	
Industry fixed effects		Yes	
Pseudo R^2		0.216	
Percent Concordant		0.832	
n		1,602	

*, **, *** Denote significance level at the 0.10, 0.05, and 0.01 level, respectively.
This table reports the relation between insider selling and the probability of firms receiving a GC opinion based on the matched sample. The detailed matching procedure is described in Section V. Panel A presents the comparison in firm characteristics between the GC firms and their matched clean firms. Panel B reports the regression results from estimating Equation (1) based on the matched sample. Wald statistics are based on the robust standard errors clustered at the client firm and auditor level.
All variables are defined in Appendix C.

Controlling for Earnings Management

Besides exerting influence on auditors, managers who have strong selling incentives may have other ways, such as earnings management, to avoid going-concern opinions. Therefore, earnings management may be a correlated omitted variable. To address this concern, we explicitly control for earnings management, where earnings management is proxied by signed performance-controlled discretionary accruals following Kothari et al. (2005), and find similar results.

Other Alternative Explanations and Alternative Measures of Insider Trading

We conduct a batch of sensitivity tests. First, to address the possibility that managers' restructuring plans drive the documented negative relation between abnormal insider sales and going-concern opinions, we include additional control variables.[24,25] Second, we apply three alternative measures of insider selling activity: (1) the net selling volumes measured as the net number of shares sold (*NSV2*); (2) the net number of selling persons (*NSP*), measured as the difference in the total number of insiders who are net sellers versus those who are net buyers in the current year, as used in prior studies (Karpoff and Lee 1991; Lee et al. 1992); and (3) the net number of sale transactions (*NSF*), as used in Huddart et al. (2007), which is the difference between the number of sale transactions and the number of purchase transactions. Third, we take into

[24] The control variables include dividend cuts, changes in capital and R&D expenditures, debt issuance, equity issuance, and restructuring activities such as spin-offs and asset sales measured in the current fiscal year. Untabulated results show that firms are less likely to receive going-concern opinions if they obtain external financing, undertake corporate restructuring, reduce dividend payout, or cut discretionary expenses. These results are consistent with prior studies (Mutchler et al. 1997; Behn et al. 2001). The results are similar when we measure restructuring activities in either year t or year t+1.

[25] In addition, we compare firms with high predicted changes in insider sales to firms with low predicted changes in insider sales. These activities, such as corporate restructure, debt or equity issuance, dividend cuts, and change in discretionary expenses, etc., show no statistical differences between the two groups of firms (untabulated).

account the effects of earnings announcements by including changes in earnings, as well as abnormal stock returns around earnings announcements, in the regressions. All reported findings are robust to these procedures. Finally, to investigate whether insiders trade opportunistically ahead of an event (the lack of going-concern opinion issuance), we test if the insider-selling pattern reverses afterward. We find that insider sales decrease in year t+1 for both the full sample and the matched sample, an indication of opportunistic insider trading in t. Furthermore, we conduct a placebo test using insider trading in t+1 as the variable of interest in Equation (1). Untabulated results show that the coefficient on the measure of insider selling becomes insignificant regardless of whether we use the predicted change or actual change in net insider selling, suggesting that our main findings likely reflect insiders' opportunistic trading behavior.

Additional Analyses

The need for a first-time going-concern report will not be apparent in many cases until late in the year. Indeed, we find that abnormal insider trading in the late period spanning (−90 days, −1 day) prior to earnings announcements has a stronger negative relation with going-concern opinions than that in the early period spanning (−365 days, −91 days) prior to earnings announcements.

Insider trading followed by an immediate bankruptcy filing will attract regulatory attention regardless of the auditor opinion. Thus, we expect our hypothesized relation to be stronger among firms with lower probability of failure. Using Altman's Z-score to measure the probability of failure, we find results consistent with this prediction. In addition, we re-run our main analysis using actual changes in net insider selling volume (*CHNSV*) and obtain similar results. Last, we re-run Equation (1) separately for surviving firms (i.e., firms without filing for bankruptcy in the subsequent two years) and bankrupt firms (i.e., firms filing for bankruptcy in the subsequent two years). This procedure is designed to identify whether the hypothesized relation is driven by auditors being less conservative (among survivors) or being too optimistic (among bankrupt firms). We find that the hypothesized relation holds for both groups with no statistical difference in this relation between the two groups.[26]

Further Discussion

SEC Rule 10b5-1, implemented in October 2000, creates a safe harbor against claims of informed trading for trades that are planned in advance. If it is common practice that insiders in distressed firms establish such plans to reduce litigation risk, particularly when they expect auditors to issue a going-concern opinion, then managers can trade safely without worrying about the audit opinion. However, the strong negative relation between insider trading and going-concern opinions that we report suggests that either establishing a 10b5-1 plan is costly for distressed firms (i.e., use of the 10b5-1 plan is not widespread among distressed firms) or that the plans are not effective in protecting insiders from legal exposure.[27]

---

[26] Further analysis indicates that, among the firms that subsequently filed for bankruptcy, the predicted changes in insider sales are much higher when they receive clean opinions, compared with firms receiving going-concern opinions. This result holds for both univariate and multivariate analyses that control for a set of variables related with future bankruptcy likelihood. Results are available upon request. Therefore, it appears that a Type II error is associated with higher levels of abnormal insider selling and that the negative association between insider sales and going-concern opinions is not merely a reflection of greater survival among firms with abnormally high insider sales. In addition, among clean firms, the levels of abnormal insider sales of firms that later filed for bankruptcy are both economically and statistically significantly higher than those of their counterparts—the survivors. We thank Professor Lynn M. LoPucki, UCLA School of Law, for providing us a detailed list of bankruptcy filings by listed companies.

[27] The SEC's recent charges against former Countrywide CEO Angelo Mozilo are a warning to corporate executives that the immunity from insider trading liability provided by 10b5-1 trading plans is not unlimited.

VI. CONCLUSION

We investigate whether insider trading affects external auditors' decisions to issue first-time going-concern opinions. Specifically, we posit that insiders have incentives to discourage auditors from issuing going-concern opinions after abnormal insider sales because the bad news from a first-time going-concern opinion is likely to attract regulators' scrutiny and class action lawsuits against insider trading from investors.

Using a sample of 12,329 financially distressed firm-year observations over the period 2000 through 2007, we find evidence that abnormal insider sales are negatively associated with the likelihood of receiving a first-time going-concern opinion. A one standard deviation increase in the predicted change in insider sales is related to a decrease of 1.39 percent in the likelihood of a going-concern opinion. Cross-sectional analyses further indicate that the economic importance of clients strengthens this negative relation, while the effect is attenuated by auditors' litigation and reputation concerns and audit committee independence. We also observe that auditors who issue clean opinions are less likely to be replaced in the following year for firms with higher levels of insider selling activity. Last, we find that the inverse relation between insider selling and going-concern opinions is weaker after SOX. The overall negative effect of insider selling on going-concern opinions is robust to the use of an alternative matched sample and three alternative measures of insider trading, to controlling for possible endogeneity and earnings management, and to the application of various methods aiming to control for correlated omitted variables. Collectively, this study highlights the association between managerial trading activity and auditor behavior.

However, we acknowledge that this study provides no direct evidence that managers openly discourage auditors from issuing going-concern opinions when engaging in insider sales. In particular, our findings might reflect the trade-offs by auditors between their own costs of issuing clean-opinion reports for high bankruptcy risk firms and the litigation costs imposed on insiders when auditors issue going-concern reports. Alternatively, trading for liquidity reasons by corporate insiders might also explain our results. Finally, our two-step regression technique might limit our ability to detect insiders' opportunistic trading behavior.

REFERENCES

Altman, E. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance* 23 (4): 589–609.

Antle, R., P. Griffen, D. Teece, and O. Willamson. 1997. *An Economic Analysis of Auditor Independence for a Multi-Client, Multi-Service Public Accounting Firm.* Report prepared for the AICPA. New York, NY: AICPA.

Baesel, J., and G. Stein. 1979. The value of information: Inferences from the profitability of insider trading. *Journal of Financial and Quantitative Analysis* 14 (3): 553–571.

Beck, J., and S. Bhagat. 1997. Shareholder litigation, share price movements, news releases, and settlement amounts. *Managerial and Decision Economics* 18 (7-8): 563–586.

Behn, B., S. Kaplan, and K., Krumwiede. 2001. Further evidence on the auditor's going-concern report: The influence of management plans. *Auditing: A Journal of Practice & Theory* 20 (March): 13–29.

Beneish, M., and E. Press. 1993. Costs of technical violation of accounting-based debt covenants. *The Accounting Review* 68 (2): 233–257.

Beneish, M., and M. Vargus. 2002. Insider trading, earnings quality, and accrual mispricing. *The Accounting Review* 77 (4): 755–791.

Beneish, M., E. Press, and M. Vargus. 2011. Insider trading and earnings management in distressed firms. *Contemporary Accounting Research* 29 (1): 191–220.

Berger, P., E. Ofek, and I. Swary. 1996. Investor valuation of the abandonment option. *Journal of Financial Economics* 42 (2): 257–287.

American Accounting Association

Bonner, S., Z.-V. Palmrose, and S. Young. 1998. Fraud type and auditor litigation: An analysis of SEC accounting and auditing enforcement releases. *The Accounting Review* 73 (4): 503–552.

Business Wire. 2000. Wolf Haldenstein Adler Freeman and Herz LLP commenced a securities fraud action against Drkoop.Com, Inc. and certain directors and officers. (July 18).

Carcello, J., and T. Neal. 2000. Audit committee composition and auditor reporting. *The Accounting Review* 75 (4): 453–467.

Carcello, J., and T. Neal. 2003. Audit committee characteristics and auditor dismissals following "new" going-concern reports. *The Accounting Review* 78 (1): 95–117.

Carcello, J., and Z.-V. Palmrose. 1994. Auditor litigation and modified reporting on bankrupt clients. *Journal of Accounting Research* (Supplement): 1–30.

Carcello, J., D. Hermanson, and F. Huss. 1995. Inappropriate audit partner behavior: Views of partners and senior managers. *Behavioral Research in Accounting* (Supplement): 245–268.

Carlson, S., G. Glezen, and M. Benefield. 1998. An investigation of investor reaction to the information content of a going concern audit report while controlling for concurrent financial statement disclosures. *Quarterly Journal of Business and Economics* 37 (3): 25–39.

Chambers, D., and J. Payne. 2008. *Audit Quality and Accrual Anomaly*. Working paper, Kennesaw State University and University of Kentucky.

Chang, H., S. Liu, and K. Wan. 2011. *The Effect of the Accelerated Filing Requirement of the Sarbanes-Oxley Act on Insider Trading*. Working paper, Drexel University, Pacific Gas and Electric Company, and University of Hong Kong.

Cheng. Q., and K. Lo. 2006. Insider trading and voluntary disclosures. *Journal of Accounting Research* 44 (5): 815–848.

Chow, C., and S. Rice. 1982. Qualified audit opinions and auditor switching. *The Accounting Review* 57 (2): 326–335.

Chung, H., and S. Kallapur. 2003. Client importance, nonaudit services, and abnormal accruals. *The Accounting Review* 78 (4): 931–955.

Craswell, A., J. Francis, and S. Taylor. 1995. Auditor brand name reputations and industry specializations. *Journal of Accounting and Economics* 20 (3): 297–322.

DeAngelo, L. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3 (3): 183–199.

DeFond, M., K. Raghunandan, and K. Subramanyam. 2002. Do non-audit service fees impair auditor independence? Evidence from going-concern audit opinions. *Journal of Accounting Research* 40 (4): 1247–1274.

Dopuch, N., R. Holthausen, and R. Leftwich. 1987. Predicting audit qualifications with financial and market variables. *The Accounting Review* 62 (3): 431–454.

Duchin, R., J. Matsusaka, and O. Ozbas. 2010. When are outsider directors effective? *Journal of Financial Economics* 96 (2): 195–214.

Dunn, K., and B. Mayhew. 2002. Auditor firm industry specialization and client disclosure quality. *Review of Accounting Studies* 9 (1): 35–58.

Fama, E., and J. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy* 81 (3): 607–636.

Finnerty, J. 1976. Insiders and market efficiency. *Journal of Finance* 31 (4): 1141–1148.

Firth, M. 1978. Qualified audit reports: Their impact on investment decisions. *The Accounting Review* 53 (3): 642–650.

Fleak, S., and E. Wilson. 1994. The incremental information content of the going concern audit opinion. *Journal of Accounting, Auditing and Finance* 9 (1): 149–166.

Francis, J., and M. Yu. 2009. The effect of Big 4 office size on audit quality. *The Accounting Review* 84 (5): 1521–1552.

Garfinkel, J. 1997. New evidence on the effects of federal regulations on insider trading: The insider trading and Securities Fraud Enforcement Act. *Journal of Corporate Finance* 3 (2): 89–111.

Givoly, T., and D. Palmon. 1985. Insider trading and the exploitation of inside information: Some empirical evidence. *Journal of Business* 58 (1): 69–87.

Guo, L., and R. Masulis. 2012. *Board Structure and Monitoring: New Evidence from CEO Turnover.* Working paper, Australian School of Business.

Hennes, K., A. Leone, and B. Miller. 2008. The importance of distinguishing errors from irregularities in restatement research: The case of restatements and CEO/CFO turnover. *The Accounting Review* 83 (6): 1487–1519.

Hoitash, R., and U. Hoitash. 2008. The role of audit committees in managing relationships with external auditors after SOX: Evidence from the USA. *Managerial Auditing Journal* 24 (4): 368–397.

Hopwood, W., J. McKeown, and J. Mutchler. 1994. Reexamination of auditor versus model accuracy within the context of the going-concern opinion decision. *Contemporary Accounting Research* 11 (2): 295–310.

Huddart, S., B. Ke, and C. Shi. 2007. Jeopardy, non-public information, and insider trading around SEC 10-K and 10-Q filings. *Journal of Accounting and Economics* 43 (1): 3–36.

Huddart, S., B. Ke, and K. Petroni. 2003. What insiders know about future earnings and how they use it: Evidence from insider trades. *Journal of Accounting and Economics* 35 (3): 315–346.

Iliev, P. 2010. The effect of SOX section 404: Costs, earnings quality and stock prices. *Journal of Finance* 65 (3): 1163–1196.

Jaffe, J. 1974. Special information and insider trading. *Journal of Business* 47 (3): 410–428.

Jagolinzer, A., and D. Roulstone. 2009. *The Effects of Insider Trading Regulation on Trade Timing, Litigation Risk, and Profitability.* Working paper, University of Colorado Boulder and The Ohio State University.

Jones, F. 1996. The information content of the auditor's going concern evaluation. *Journal of Accounting and Public Policy* 15 (1): 1–27.

Karpoff, J., and D. Lee. 1991. Insider trading before new issue announcement. *Financial Management* 20 (1): 18–26.

Kothari, S. P., A. Leone, and C. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1): 163–197.

Krishnan, J., L. Su, and Y. Zhang. 2011. Nonaudit services and earnings management in the pre-SOX and post-SOX eras. *Auditing: A Journal of Practice & Theory* 30 (3): 103–123.

Lee, D., W. Mikkelson, and M. Partch. 1992. Managers' trading around stock repurchases. *Journal of Finance* 47 (5): 1947–1961.

Lee, I., M. Lemmon, Y. Li, and J. Sequeria. 2011. *The Effect of Regulation on the Volume, Timing and Profitability of Insider Trading.* Working paper, KAIST Business School.

Lennox, C. 2000. Do companies successfully engage in opinion-shopping? Evidence from the UK. *Journal of Accounting and Economics* 29 (3): 321–337.

Li, C. 2009. Does client importance affect auditor independence at the office level? Empirical evidence from going-concern opinions. *Contemporary Accounting Research* 26 (1): 201–230.

Li, C., L. Sun, and M. Ettredge. 2010. Financial executive qualifications, financial executive turnover, and adverse SOX 404 opinions. *Journal of Accounting and Economics* 50 (1): 93–110.

Lim, C., and H. Tan. 2008. Non-audit service fees and audit quality: The impact of auditor specialization. *Journal of Accounting Research* 46 (1): 199–246.

Ma, Y. 2001. Insider trading behavior prior to Chapter 11 bankruptcy announcements. *Journal of Business Research* 54 (1): 63–70.

Menon, K., and D. Williams. 2010. Investor reaction to going concern audit reports. *The Accounting Review* 85 (6): 2075–2105.

Mutchler, J., W. Hopwood, and J. McKeown. 1997. The influence of contrary information and mitigating factors in audit opinion decisions on bankrupt companies. *Journal of Accounting Research* 35 (2): 295–310.

Myers, L., J. Schmidt, and M. Wilkins. 2011. *An Investigation of Recent Changes in Going Concern Reporting Decisions Among Big N and Non-Big N Auditors.* Working paper, University of Arkansas, The University of Texas at Austin, and Texas A&M University.

Nasser, T., and B. Gup. 2008. Insider trading and large Chapter 11 bankruptcies in USA. *International Journal of Banking and Finance* 5 (1): 1–29.

Noe, C. 1999. Voluntary disclosures and insider transactions. *Journal of Accounting and Economics* 27 (3): 305–326.
Palmrose, Z.-V. 1988. An analysis of auditor litigation and audit service quality. *The Accounting Review* 63 (1): 55–73.
Park, S., H. Jang, and M. Loeb. 1995. Insider trading activity surrounding annual earnings announcements. *Journal of Business Finance and Accounting* 22 (4): 587–614.
Penman, S. 1985. A comparison of the information content of insider trading and management earnings forecasts. *Journal of Financial and Quantitative Analysis* 20 (1): 1–17.
Raghunandan, K., and D. Rama. 1995. Audit reports for companies in financial distress: Before and after SAS No. 59. *Auditing: A Journal of Practice & Theory* 14 (1): 50–63.
Rashkover, B., and C. Winter. 2005. The impact of Sarbanes-Oxley on SEC enforcement in public company disclosure cases. *International Journal of Disclosure and Governance* 2: 312–324.
Reynolds, K., and J. Francis. 2000. Does size matter? The influence of large clients on office-level auditor reporting decisions. *Journal of Accounting and Economics* 30 (3): 375–400.
Richardson, S., S. Teoh, and P. Wysocki. 2004. The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives. *Contemporary Accounting Research* 21 (4): 885–924.
Rogers, J. 2008. Disclosure quality and management trading incentives. *Journal of Accounting Research* 46 (5): 1265–1296.
Roulstone, D. 2008. *Insider Trading and the Information Content of Earnings Announcements*. Working paper, The Ohio State University.
Rozeff, M., and M. Zaman. 1988. Market efficiency and insider trading: New evidence. *Journal of Business* 61 (1): 25–44.
Seyhun, H. 1986. Insiders' profits, costs of trading, and market efficiency. *Journal of Financial Economics* 16 (2): 189–212.
Seyhun, H. 1992. Effectiveness of insider trading sanctions. *The Journal of Law and Economics* 35 (1): 149–182.
Seyhun, H. 1998. *Investment Intelligence from Insider Trading*. Cambridge, MA: MIT Press.
Seyhun, H., and M. Bradley. 1997. Corporate bankruptcy and insider trading. *Journal of Business* 70 (2): 189–216.
Shu, S. 2000. Auditor resignations: Clientele effects and legal liability. *Journal of Accounting and Economics* 29 (2): 173–205.
Stanley, J., F. Zoort, and G. Taylor. 2009. The association between insider trading surrounding going concern audit opinions and future bankruptcy. *Management Auditing Journal* 24 (3): 290–312.
Watts, R., and J. Zimmerman. 1981. Auditor Independence. Unpublished paper, University of Rochester.
Weil, J. 2001. Did accountants fail to flag problems at dot-com casualties? *Wall Street Journal* (February 9).

APPENDIX A
FIRST-STAGE REGRESSION

This appendix describes the first-stage model used to estimate *PREDICT_CHNSV*, reports the summary statistics for variables in the model, and presents the results of the estimation.

Panel A: First-Stage Regression Model

$$\begin{aligned} CHNSV_t = {} & \alpha_0 + \alpha_1(OPTION\_GRANTS_{t-1}) + \alpha_2(RETURN_{t-1}) + \alpha_3(CHNSV_{t-1}) \\ & + \alpha_4(\Delta MB_{t-1}) + \alpha_5(\Delta MV_{t-1}) + \alpha_6(\Delta ROE_{t-1}) + \varepsilon_t. \end{aligned} \tag{A1}$$

CHNSV = difference between the current year's net selling volume (*NSV*) and the average *NSV* during the prior four years, where *NSV* is measured as the natural logarithm of 1 plus the absolute total dollar amount of net insider trading (in thousands) over a fiscal year, with a positive (negative) sign added for net sales (net purchases). Net insider trading is insider purchases subtracted from insider sales;

OPTION_GRANTS = number of options granted, scaled by the number of outstanding shares. Following Cheng and Lo (2006), we use the resulting options held by insiders from the derivatives trading data in the insider trading database;

RETURN = firm's annual stock returns;

ΔMB = difference between the current year's market-to-book ratio and the average market-to-book ratio during prior years up to four, where the market-to-book ratio is measured as the market value of equity to the book value of equity. The market value of equity is the product of the year-end stock price and the number of outstanding common stocks (Compustat item PRCC_F * CSHO). The book value is Compustat item CEQ;

ΔMV = difference between the current year's firm size and the average firm size during prior years up to four, where firm size is measured as the natural logarithm of the market value of equity defined above (Compustat item PRCC_F * CSHO); and

ΔROE = difference between the current year's return on equity (*ROE*) and the average *ROE* during prior years up to four, where *ROE* is measured as the bottom line earnings scaled by the book value of equity (Compustat item NI/CEQ).

We estimate Equation (A1) using all firm-year observations in the intersection of the insider trading dataset and the Compustat database, rather than the financially distressed sample, to increase the precision of estimating the expected insider trading. Our results are robust to the use of the financially distressed sample in the first step of the regression. As in Cheng and Lo (2006), the estimation is performed annually over the sample period to mitigate the time effect (Fama and MacBeth 1973).

Panel B: Summary Statistics for Variables in the First-Stage Regression

Variable	n	Mean	Median	Std. Dev.	Q1	Q3
$OPTION\_GRANTS_{t-1}$	49,820	0.006	0.001	0.015	0.000	0.004
$RETURN_{t-1}$	49,820	0.065	0.054	0.843	−0.212	0.251
$CHNSV_{t-1}$	49,820	0.288	0.000	4.209	−1.722	5.922
MB_{t-1}	49,820	2.970	1.881	4.623	1.163	3.296
MV_{t-1}	49,820	5.798	5.724	2.143	4.121	7.221
ROE_{t-1}	49,820	−0.052	0.079	0.744	−0.053	0.147

Panel C: Results for the First-Stage Regression

Variables	Coeff. Estimate		t-value
Intercept	0.207		1.49
$OPTION\_GRANTS_{t-1}$	8.633***		4.55
$RETURN_{t-1}$	0.489***		3.75
$CHNSV_{t-1}$	0.195***		27.27
ΔMB_{t-1}	0.028**		3.02
ΔMV_{t-1}	0.285***		4.36
ΔROE_{t-1}	0.071**		3.05
Avg. Adj. R^2		0.161	
Observations		49,820	

*, **, *** Denote significance at the 0.10, 0.05, and 0.01 levels, respectively.

APPENDIX B

MOTIVATION FOR, AND EXPECTATIONS OF, THE EXPLANATORY VARIABLES IN THE GOING-CONCERN MODEL

The motivation for the choice of explanatory variables used in the going-concern model (Equation (1)) is based on prior studies regarding the determinants of going-concern opinions (e.g., Dopuch et al. 1987; Mutchler et al. 1997; Reynolds and Francis 2000; DeFond et al. 2002). *ZSCORE* is the Altman's Z-score, with higher values indicating a lower probability of bankruptcy. We also include *LOSS* (a dummy indicating a loss in the current year) because firms with losses are more likely to fail (Reynolds and Francis 2000). The natural logarithm of total assets (*SIZE*) is included because large firms have more negotiating power in the event of financial difficulties and, hence, are more likely to avoid bankruptcy (Reynolds and Francis 2000). The *AGE* variable is the natural logarithm of the number of years a company has been publicly traded. Younger firms are more prone to failure and, therefore, have a higher likelihood of receiving a going-concern report (Dopuch et al. 1987).

We also include two market-based measures following Dopuch et al. (1987): *RETURN*, which is the stock return over a fiscal year, and *VOLATILITY*, which is the standard deviation of the company's monthly stock returns. *RETURN* is expected to be negatively associated with *GC OPINION* and *VOLATILITY* to be positively associated with *GC OPINION*. Other determining factors in our model include leverage (*LEV*) and change in leverage (*CLEV*), because Mutchler et al. (1997) find that debt covenant violations are positively associated with the probability of a going-concern opinion being issued. We include *LEV* to capture proximity to covenant violation as firms close to violation are likely to have high leverage (Beneish and Press 1993), and we include *CLEV* because increases in leverage are likely to move firms closer to covenant violation (Reynolds and Francis 2000). *OCF* (operating cash flows divided by total assets) is included because poor operating cash flows are often associated with the probability of bankruptcy and because the Altman's Z-score does not include a cash-flow measure. In addition, we include *ANNLAG* (the number of days between the fiscal year-end and the earnings announcement date) because prior research finds that going-concern opinions are associated with longer reporting delays (Raghunandan and Rama 1995; Carcello et al. 1995).

In addition, several factors are included that are likely to reduce the probability of receiving a going-concern opinion. *INVESTMENTS* is the sum of a firm's cash and investment securities (short- and long-term), scaled by total assets. Firms with large cash and investment securities have a greater ability of avoiding bankruptcy in the event of financial difficulty (Berger et al. 1996). *NEWFINANCE* captures new financing activities, mainly the new issuance of debt or equity securities (public or private) over the succeeding fiscal year. Mutchler et al. (1997) find that new financing reduces the probability of bankruptcy. The information about new financing activities is based on the reported numbers in the statement of cash flows (Compustat item DLTIS and SSTK). Specifically, it takes the value of 1 if a firm has cash flows from the issuance of debts (i.e., DLTIS > 0) or the proceeds from the issuance of equity securities are over 5 percent of the firm's market value of equity (i.e., SSTK > 5 percent of the firm's market value). We follow Richardson et al. (2004) and require proceeds from equity issuance to be greater than 5 percent of the market value of equity for that year to filter out employees' exercise of stock options. We also include *BIGN*, an indicator variable equal to 1 if the auditor is a member of the Big N auditing firms, and 0 otherwise. This variable is included because prior research argues that Big N auditors are more likely to issue going-concern audit opinions (Mutchler et al. 1997).

APPENDIX C
DEFINITION OF VARIABLES

Variables in the Model of Going-Concern Opinions

GC OPINION = indicator equal to 1 if the firm receives a first-time going-concern opinion, and 0 otherwise;
PREDICT_CHNSV = predicted value of the change in net insider selling volume (*CHNSV*) based on the model of insider trades in Appendix A;
CHNSV = difference between the current year's net selling volume (*NSV*) and the average *NSV* during the prior four years, where *NSV* is measured as the natural logarithm of 1 plus the absolute total dollar amount of net insider trading (in thousands) over a fiscal year, with a positive (negative) sign added for net sales (net purchases). Net insider trading is insider purchases subtracted from insider sales;
ZSCORE = Altman's (1968) Z-score;
LOSS = indicator equal to 1 if the client's net income (Compustat item NI) is negative for the current year, and 0 otherwise;
SIZE = natural logarithm of total assets (Compustat item AT);
AGE = natural logarithm of the number of years of data for the client firm since the coverage in Compustat;
RETURN = firm's cumulative stock return over the current year;
VOLATILITY = standard deviation of monthly stock returns over the current year;
LEV = ratio of total liabilities (Compustat item LT) to total assets at the end of the current year;
CLEV = change in *LEV* from previous year to current year;
OCF = operating cash flow (Compustat item OANCF) scaled by total assets for the current year;
ANNLAG = number of days between the fiscal year-end and earnings announcement date for the current year;
INVESTMENT = short- and long-term investment securities (including cash and cash equivalents) (Compustat items CHE and IVPT), scaled by total assets;
NEWFINANCE = indicator variable equal to 1 if the client has a new issuance of equity or debt over the subsequent fiscal year (i.e., nonzero Compustat item DLTIS or the amount of Compustat item SSTK is over 5 percent of the firm's market value of equity), and 0 otherwise; and
BIGN = indicator equal to 1 if the auditor is a member of the Big 5 before 2002 or a member of the Big 4 after 2002, and 0 otherwise.

Variables in the Models of Cross-Sectional Tests

DEPENDENCE = ratio of a specific client's audit fee (market value) to the total audit fees (total market values) for all clients of an incumbent auditor at the local office, where the fees and clients' market values are in natural logarithms;
FRAUD = 1 for clients if any of the local auditor office's clients announced a financial fraud in the recent two years, and 0 otherwise;
EXPERT = 1 if the incumbent auditor ranks as a top two firm in the client's industry (two-digit SIC code) in terms of market share of sales, and 0 otherwise;
OFFICESIZE = natural logarithm of the total audit fees of a local practice office in a given fiscal year; and
INDEP = 1 if the firm has a fully independent audit committee (i.e., all the audit committee members are independent), and 0 otherwise.

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 395–427

American Accounting Association
DOI: 10.2308/accr-50320

Economic Consequences of Mandated Accounting Disclosures: Evidence from Pension Accounting Standards

Elizabeth C. Chuk
University of Southern California

ABSTRACT: I examine whether firms alter their behavior in response to changes in accounting standards that mandate new financial statement disclosures. While prior research suggests that new *recognition* rules lead to changes in firm behavior, there is limited evidence that *disclosure* rules can impact firm behavior. This study helps to fill this void in the literature by examining the economic consequences of the mandated disclosures of pension asset composition required under SFAS 132R. Under pension accounting rules, the composition of pension assets is a key determinant of the assumed expected rate of return (ERR) on pension assets. I find that when firms disclose asset composition for the first time under SFAS 132R, firms that were previously using upward-biased ERRs respond by increasing asset allocation to high-risk securities and/or reducing the ERR assumption. While disclosure requirements arguably create less powerful incentives to alter firm decisions than recognition requirements, these findings offer evidence that firms alter behavior in response to disclosure standards.

Keywords: *economic consequences of accounting standards; disclosure; pensions; SFAS 132R.*

Data Availability: *The data used in this study are publicly available from the sources indicated in the text.*

I thank Dave Burgstahler (dissertation chairman) for continuous feedback and guidance in the development of this paper. I also thank other members of my dissertation committee—Richard Hartman, Jon Karpoff, Dawn Matsumoto, Shiva Rajgopal, and Terry Shevlin—as well as Linda Bamber, Brad Blaylock, Andy Call, Joe Comprix, Mark DeFond, Alex Edwards, Fabrizio Ferri, Jenny Gaver, Michelle Hanlon, Trevor Harris, Urooj Khan, Stacie Laplante, Jim Naughton, Marc Picconi, D. Shores, Mark Soliman, Mauricio Soto, KR Subramanyam, Lloyd Tanlu, Bob Trezevant, Aida Wahid, Ryan Wilson, Jenny Zhang, and seminar participants at Columbia University; New York University; Rice University; and the University of British Columbia; University of California, Berkeley; The University of Georgia; Southern California University; The University of Utah; and University of Washington. I also thank Steven Kachelmeier (former senior editor) and two anonymous referees for their helpful comments. I gratefully acknowledge financial support from the Deloitte Foundation.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: January 2011
Accepted: October 2012
Published Online: October 2012

I. INTRODUCTION

Prior research suggests that changes in accounting recognition standards lead firms to alter some types of transactions (Mittelstaedt et al. 1995; Graham et al. 2005; Bens and Monahan 2008; Choudhary et al. 2008; Zhang 2009; Amir et al. 2010). However, these prior studies focus on accounting standards mandating *recognition*, as opposed to *disclosure*. Disclosure requirements arguably create less powerful incentives to alter firm decisions than recognition requirements. The purpose of my study is to examine whether firms alter their behavior in response to accounting standards that mandate new financial statement disclosures.

My study empirically examines this question using defined benefit pension accounting as a research setting.[1] In particular, I examine whether firms alter behavior in response to the mandated disclosure of pension asset composition under SFAS 132R (FASB 2003). The content of the SFAS 132R disclosure requirement is economically important because it relates to an assumption affecting net income. The expected rate of return (ERR) assumption determines the expected return on pension assets and, all else equal, a higher assumed ERR results in higher reported earnings. Although SFAS 132R only changes disclosure requirements for the allocation of pension assets among various investment categories, this allocation is a key determinant of the ERR. Therefore, I predict that the disclosure of the asset allocation constrains the ability to manage earnings using the ERR.

Because of its importance, the ERR has always been subject to scrutiny. During the early 2000s, some financial analysts and regulators speculated that firms' ERR assumptions were unrealistically high. In 2002, the SEC publicly warned companies that it might challenge ERRs above 9 percent (*Bloomberg Businessweek Magazine* 2004). Then in 2003, the FASB issued SFAS 132R to require disclosure of the percentage composition of pension assets among major investment categories.

I hypothesize that firms previously using upward-biased ERRs respond to SFAS 132R by a combination of increasing the proportion of pension assets allocated to high-risk investments to justify the inflated ERR and reducing the ERR assumption to reflect the true riskiness of the pension assets. I predict that the magnitudes of these responses depend on the extent to which the pre-SFAS 132R ERR is higher than justified by pension asset composition.

I begin by constructing a proxy to measure the extent to which the pre-SFAS 132R ERR is higher than that justified by pension asset composition. My proxy is derived from a model that estimates firms' ERRs as a function of their reported pension asset allocation. Specifically, I regress the ERR on the proportions of pension assets invested in various categories of securities. I assume that SFAS 132R reduces the opportunity for firms to report ERRs that are not justified by their asset composition, such that there is greater alignment between the ERR and asset composition in the period after SFAS 132R. Based on this assumption, I use all firm-years after SFAS 132R to obtain parameter estimates that describe the unbiased relation between the ERR and asset composition. Using these parameter estimates, I then compute prediction errors in the year before SFAS 132R for

[1] Unless otherwise stated, all references to pension plans hereafter refer to defined benefit plans only. In a defined benefit (DB) plan, the employee's pension benefit is determined by a formula that takes into account the years of service, salary, age at retirement, and other factors. In contrast, in a defined contribution (DC) plan, contributions are paid into an individual account for each participant, and contributions are invested and credited to the individual's account. The employer (employee) bears the investment risk under DB (DC) plans. DB pensions in the U.S. have $2 trillion in aggregate pension assets under management, and in 2008 approximately 25 percent of firms on Compustat had DB pensions. Firms with DB plans are economically significant as the aggregate market capitalization of firms with DB pensions represents 68 percent of the total market capitalization of all firms on Compustat.

each firm as a measure of the portion of the ERR that is not explained by pension asset allocation. For simplicity, I hereafter refer to this measure as the unexplained ERR.

My research design uses cross-sectional regressions, where the changes in firm behavior are measured in the fiscal years immediately before and after the issuance of SFAS 132R. Because the two firm responses of interest are potentially simultaneously determined, I use two-stage least squares, where the dependent variables are the change in ERR and the change in asset allocation, and the independent variable of interest is the unexplained ERR.

Consistent with my predictions, I find that both firm responses are associated with the unexplained ERR. In particular, when disclosure of asset composition in financial statements is required for the first time under SFAS 132R, firms whose pre-SFAS 132R ERRs are higher than justified by their pension asset composition respond to SFAS 132R by increasing their asset allocation to high-risk securities and/or reducing their ERRs. Furthermore, I find that the magnitudes of both responses are positively associated with the upward bias in the pre-SFAS 132R ERR.

These findings make an important contribution to the literature on managerial responses to accounting standards. Specifically, the results from this study are consistent with the idea that firms are willing to undertake "real" actions in order to report favorable accounting information (Graham et al. 2005). My study complements prior studies that examine whether firms alter their behavior in response to changes in accounting standards related to recognition rules (Mittelstaedt et al. 1995; Bens and Monahan 2008; Choudhary et al. 2008; Zhang 2009; Amir et al. 2010) by extending the research question to disclosure rules.

The distinction between recognition and disclosure is important. Prior accounting research has documented differences between amounts recognized in financial statements and amounts disclosed in footnotes with respect to several dimensions, such as value relevance (Barth 1991; Hann et al. 2007) and the level of auditor tolerance for misstatements (Libby et al. 2006). In addition, Amir and Ziv (1997) model a firm's decision between recognition and disclosure under early adoption of SFAS 106 and find evidence that firms do not view recognition and disclosure as equivalent because of contracting effects. Therefore, while prior studies find that firms alter their behavior in response to changes in *recognition* rules, these prior studies do not address the question of whether firms alter their behavior in response to changes in *disclosure* rules.

SFAS 132R offers a natural experiment for studying the effects of accounting standards on firm behavior. SFAS 132R was largely unanticipated by firms (as discussed in Section II), such that the effects of the rule change are likely concentrated in the year of the rule change. In the first year of SFAS 132R, firms are required to disclose historical pension asset allocations in addition to current-year information. Because firms did not anticipate SFAS 132R, the historical asset allocations represent the allocations chosen by firms in the absence of mandated disclosure of pension asset composition. I exploit the features of SFAS 132R to construct a control sample that is not subject to the mandated disclosure of asset allocation.

In robustness tests, I examine a potential alternative explanation. As with any test of changes over time, a limitation to my study is that there may be potential confounding events concurrent with SFAS 132R. Therefore, it is possible that the firm responses observed in my study are driven by events other than the issuance of SFAS 132R. Specifically, I examine whether the observed managerial responses are driven by a warning issued by the SEC that it might challenge unreasonable pension assumptions. The results of the robustness tests support the main findings that the firm responses are attributable to SFAS 132R.

Section II motivates the study and describes the relevant FASB accounting standards for DB pensions. Section III develops hypotheses. Section IV describes the data and empirical methods. Section V discusses results and robustness tests. Section VI concludes.

II. MOTIVATION

Firm Responses to Accounting Rules

Several prior studies provide evidence that firms alter some types of transactions in response to changes in accounting standards. Zhang (2009) finds that firms reduce their speculative use of derivative instruments after SFAS 133 required recognition of all derivatives as either assets or liabilities at fair value. Choudhary et al. (2008) document that firms accelerate the vesting of employee stock options to avoid recognizing unvested option grants at fair value under SFAS 123R. Bens and Monahan (2008) find that banks reduce their investments in asset-backed commercial paper and enter into costly restructuring arrangements to avoid off-balance sheet consolidation under FIN 46. Mittelstaedt et al. (1995) document that the reductions in retiree health care benefits are associated with the adoption of SFAS 106, which requires the recognition of underfunded health care liabilities. Imhoff and Thomas (1988) find a substitution from capital leases to operating leases after SFAS 13 required the recognition of capital leases on the balance sheet. Beatty (1995) documents that firms that early-adopt SFAS 115 alter their investment behavior. Horwitz and Kolodny (1980) find a reduction in R&D spending after SFAS 2 required R&D to be expensed. Prior literature has also examined firm responses to accounting standards related to pensions. Amir et al. (2010) find that firms shift pension assets from equities to bonds after the required balance sheet recognition of the pension-funding status under SFAS 158 and FRS 17 in the U.S. and U.K., respectively.[2]

My study differs from these prior studies, primarily by examining a setting in which only *disclosure* requirements change. Disclosure requirements arguably create less powerful incentives to alter firm decisions than recognition requirements.[3] A finding that changes in disclosure requirements alone are enough to induce meaningful changes in firm behavior would offer compelling evidence in support of the idea that firms do alter their behavior in response to changes in accounting standards. My study contributes to the literature on the effects of accounting regulation by extending the research question to accounting standards that only mandate disclosure.

Pension Accounting as a Research Setting

SFAS 132R, *Employers' Disclosures about Pensions and Other Postretirement Benefits*, offers a unique natural experiment with several advantages for studying the effects of accounting standards on firm behavior. First, the consequences resulting from the rule change are likely concentrated and observable in the year that the accounting standard was issued. SFAS 132R became effective quickly because the FASB believed that most of the disclosure requirements would be "readily available in firms' existing information systems" (SFAS 132R, ¶A28). The FASB added a project on pension disclosures to its technical agenda in March 2003, issued an Exposure Draft in September 2003, and

---

[2] When FRS 17 was issued in the U.K., it was a disclosure standard for three years before becoming a recognition standard. However, the eventual transition from disclosure to recognition was one of FRS 17's stated goals at the time of issuance. FRS 17 "allows for a long implementation period, with disclosures building up in the notes to the accounts" (FRS 17, ASB 2000, ¶57 in Appendix IV). At issuance, firms were aware that FRS 17 would eventually require recognition of the funding status. Therefore, the economic consequences documented by Amir et al. (2010) could be due to firms' response to the recognition rule, as opposed to the disclosure rule. My research setting using SFAS 132R pertains to a standard that alters disclosure only.

[3] I assume that auditors check for the reasonableness of the ERR prior to SFAS 132R but more rigorously test whether the ERR is supported by the asset allocation after SFAS 132R because of higher litigation risk for information actually published in the 10-K. This assumption is consistent with Libby et al. (2006), who find that auditors permit more misstatements in disclosed, as opposed to recognized, amounts. Extending this logic, I expect that auditors permit more misstatements in amounts that do not even appear in financial statements as opposed to disclosed amounts.

the disclosure requirements became effective for fiscal years ending after December 15, 2003. Thus, firms likely did not anticipate SFAS 132R's new required disclosures before the final issuance of the standard.[4] Second, SFAS 132R mandates that the "disclosures for earlier annual periods presented for comparative purposes should be restated for the percentages of each major category of plan assets held" (SFAS 132R, page 3 of Summary), such that in the first year that SFAS 132R was effective, prior-year asset composition was required to be disclosed in addition to current-year asset composition. Because firms likely did not anticipate SFAS 132R, they likely did not expect that the prior-year asset allocations would be disclosed in the financial statements. Thus, the historical asset allocations represent asset allocations chosen by firms in the absence of mandated disclosure of pension asset composition to justify their ERRs. In my empirical design, I exploit these features of SFAS 132R and use the historical asset allocations to construct a control sample that is not subject to the mandated disclosure rules. Third, SFAS 132R only alters disclosure requirements without any contemporaneous changes to recognition or measurement rules. The absence of such changes reduces the potential confounding effects that are attributable to recognition or measurement instead of disclosure. Fourth, pensions offer an interesting setting with large and economically meaningful effects to study the consequences of accounting standards.

As in prior studies examining the economic consequences of accounting standards, SFAS 132R is not necessarily the sole cause of changes in managerial behavior related to pensions. It is important to consider other factors that can concurrently affect changes in managerial behavior. For instance, actual pension asset returns and the overall equity market performance can contribute to the observed changes in ERR and/or asset allocation, even in the absence of managerial responses to SFAS 132R. As discussed in Section IV, I empirically control for these other factors to examine whether the impact of SFAS 132R on managerial behavior is incremental to other factors that can potentially influence firm decisions related to pensions.

Institutional Background for Pension Accounting (SFAS 132R and SFAS 87)

SFAS 132R, *Employers' Disclosures about Pensions and Other Postretirement Benefits*, provides disclosure guidelines for the plan assets and obligations for firms with pension plans. SFAS 132R became effective for fiscal years ending after December 15, 2003 and revises the disclosures originally mandated by SFAS 132, which took effect for fiscal years beginning after December 15, 1997. The revised standard requires annual disclosure of the percentage composition of major categories of pension plan asset allocation (i.e., equities, bonds, real estate, and other), along with narrative descriptions of investment strategies and the basis used to determine the overall expected long-term rate of return on assets.[5]

---

[4] I am not able to rule out the possibility that firms anticipated SFAS 132R. There are two potential implications. First, if firms anticipated SFAS 132R and adjusted their pensions accordingly some time during 2003, then the economic consequences documented in my study are partially due to the anticipation of the new mandate, as opposed to the passage of the mandate. For my research question, both types of economic consequences are important. Second, if firms adjusted their pensions in 2002 or before, then there would be measurement error in my proxies for economic consequences, thereby reducing the power of my tests.

[5] SFAS 132R includes an example of pension asset allocation using these four broad categories, and firms typically interpret the examples provided by FASB as required disclosures and use them as templates when preparing their financial statements (Zion and Carache 2005). Also, SFAS 132R requires other new annual disclosures, which are not the focus of my study, including: accumulated benefit obligation (this disclosure had been eliminated under the original SFAS 132); target allocations of each plan asset category for companies that use target allocations; benefits expected to be paid in each of the next five fiscal years and in the aggregate for the five fiscal years thereafter; the best estimate of contributions expected to be paid to the plan during the next fiscal year; separate statement of the assumptions used to determine the benefit obligation and the net benefit cost; and the measurement dates used to determine assets and benefit obligations. SFAS 132R also requires disclosure of the separate components of net pension cost and estimated and actual contributions to pension plans in interim financial reports.

SFAS 132R only enhances disclosure and does not change the measurement or recognition rules under SFAS 87, *Employers' Accounting for Pensions* (FASB 1985). The pension cost computed under SFAS 87 includes three main components: (1) service cost, (2) interest cost, and (3) expected return on pension plan assets, which is the product of the ERR assumed by management and the fair value of pension plan assets.[6] To mitigate concerns about the volatility of returns on pension assets, SFAS 87 allows firms to use expected returns, rather than actual returns, and differences between expected and actual returns are amortized over time using the "corridor" approach.[7]

III. HYPOTHESIS DEVELOPMENT

According to pension accounting rules, the expected return on pension assets offsets the service cost and interest cost when computing pension expense, such that a higher ERR assumption results in higher reported earnings. Changes in the ERR can result in economically significant changes in reported earnings. Zion and Carache (2002) estimate that aggregate profits for the S&P 500 firms with DB plans would decline by an estimated $44 billion in 2003 alone if these firms lowered their ERR to 6.5 percent.[8]

Prior research has examined the use of pension assumptions to manage earnings (Bergstresser et al. 2006), and using the ERR to manage earnings appears to be effective because the market does not fully impound pension information (Picconi 2006; Franzoni and Marin 2006). In the years preceding SFAS 132R, some analysts and regulators speculated that ERRs were too high when compared to historical returns in equity and bond markets. For instance, a J.P. Morgan Asset Management investment strategist argued in a 2002 *Fortune* article that a far more reasonable assumption for returns would be between 6.5 percent and 7 percent, given the widely held expectations for stock and bond performance for the next 20 years (Revell 2002). In December 2002, the SEC publicly warned companies that it might challenge rate-of-return assumptions above 9 percent. In October 2004, the SEC launched an investigation into six companies regarding their pension assumptions. "Among other things, we're looking to see if companies have reverse-engineered the rates to get to a certain financial result," said Kenneth Lench, an assistant director in the SEC's enforcement division (Stuart 2005).[9]

SFAS 132R requires annual disclosure of the percentage composition of major categories of pension plan assets, which allows users to assess the reasonableness of the ERR with respect to the riskiness of the pension assets' investment categories. Bergstresser et al. (2006) find that managers use asset allocations to justify their ERRs, but Amir and Benartzi (1998) find that the ERR is only

[6] Other components include: the amortization of unrecognized prior service costs, which represent retroactive benefits earned prior to the year of initiating or amending a DB plan; the amortization of gains or losses from the effects of changing actuarial assumptions and differences between actual and expected return on plan assets; and the amortization of unrecognized net obligation or asset existing before the adoption of SFAS 87.

[7] Under the corridor approach, the cumulative unrecognized net gain or loss in excess of a "corridor" defined as 10 percent of the greater of the projected benefit obligation or the market value of plan assets is the minimum amount amortized.

[8] The most controversial and contentious assumption is the ERR (Zion and Carache 2005). I focus on the ERR as the major assumption available for managers to manipulate pension expense, although there are other important pension assumptions, including the discount rate, salary inflation rate, mortality, employee turnover, and retirement age.

[9] The six companies were Boeing, Delphi, Ford, General Motors, Navistar, and Northwest Airlines. Although this sample is too small to conduct any formal statistical tests, the business press has indicated that these firms have suffered losses in market value as a result of the SEC investigation. In fact, on the day that General Motors publicly acknowledged the SEC investigation, the company's stock price declined by 7 percent over the course of one day (Bruno 2005).

"weakly" related to asset allocation. Both of these studies use pre-SFAS 132R samples. I predict that the mandated disclosure of asset composition constrains the opportunity for firms to report ERRs that are not justified by actual asset allocations. If firms report upward-biased ERRs prior to SFAS 132R, then I predict that they reduce the bias after SFAS 132R, thereby strengthening the relation between ERR and asset composition.

H1: Firms that use inflated ERRs prior to SFAS 132R are more likely to increase their asset allocation to higher-risk securities after SFAS 132R.

H2: Firms that use inflated ERRs prior to SFAS 132R are more likely to decrease their ERRs after SFAS 132R.

My maintained assumption is that opportunistic firms react to the new FASB requirement to disclose pension asset allocation in the 10-K, despite the fact that some firms were already disclosing asset composition prior to SFAS 132R. There are three alternative sources of pension asset allocation information prior to SFAS 132R.

The first source comes directly from firms' annual reports. Prior to SFAS 132R, SFAS 87 acknowledged the importance of asset allocation disclosures (SFAS 87, ¶54), but did not enforce the disclosures. As a result, some firms voluntarily disclosed asset allocation. Amir and Benartzi (1998) examine voluntary pension asset allocation disclosures under SFAS 87 for a sample of 27 firms. In their sample (Amir and Benartzi 1998, Appendix), most firms either do not disclose pension asset allocation at all or provide only a qualitative list of the general asset categories. Only two firms in their sample disclose the percentage composition of pension assets, which is the format mandated under SFAS 132R.

The second source of asset allocation information is IRS Form 5500, which is a required filing under ERISA for pension plans with more than 100 participants. While Form 5500 contains a larger list of finer asset categories than under SFAS 132R, firms report a majority of their assets on Form 5500 in categories labeled as "common/collective trusts" or "master trust investment accounts." For instance, 56 percent of DB assets for all firms that filed Form 5500 in 2007 were collectively classified as assets held in common/collective trusts or master trusts.[10] Another issue is that a firm may have pension plans operating under different tax filer identifiers, making it difficult for analysts and investors to aggregate all the plans for a given firm, as under SFAS 132R. Furthermore, there is a two-year lag between the filing date and the time when the forms become publicly available (Jin et al. 2006). The lag time reduces the likelihood that managers feel pressure to have asset compositions that support the ERR. For these reasons, Form 5500 is unlikely to be a direct substitute for the SFAS 132R disclosures. Nevertheless, I reperform my main tests using Form 5500 data; the untabulated results (discussed in Section V) support the idea that SFAS 132R provides incremental information to Form 5500.

The third source is proprietary databases, through which some firms voluntarily disclose pension asset composition. One of the most prominent databases is compiled by Pensions & Investments (P&I). Approximately 600 of the 1,000 plans covered by P&I are corporate pensions, representing less than half of the 1,300 firms on average per year in Compustat that have pensions during my sample period. There are no assurances that firms disclose truthfully to the proprietary databases. Furthermore, similar to IRS Form 5500, P&I is also not a timely source; P&I typically

---

[10] Source: http://www.dol.gov/ebsa/PDF/2007pensionplanbulletin.PDF. Also, to offer a concrete example, in 2007 General Motors reported the following composition of its $29.96 billion in DB pension assets on Form 5500: $29.93 billion (99.9 percent) in "common/collective trusts" and the remaining $0.03 billion (0.1 percent) in "other."

has a one-year lag.[11] In summary, while these three sources provide important information prior to SFAS 132R, it is likely that SFAS 132R provides incremental information that allows users to better evaluate the validity of the ERR.

It is possible that the disclosures available prior to SFAS 132R provide sufficiently strong incentives for firms to report ERRs consistent with asset allocation. Under this scenario, the ERR and asset composition would already be aligned prior to SFAS 132R, and there would not be significant changes in firm behavior when asset composition is disclosed for the first time in financial statements (i.e., results would not support H1 and H2).

IV. DATA AND EMPIRICAL DESIGN

My sample consists of the Compustat universe of U.S. firms with DB pensions. More specifically, I require nonmissing ERR, actual returns to pension assets, pension assets, pension liabilities, and pension asset allocation data. I further impose the criterion that the pension asset allocation percentages sum to 100 percent, after I manually adjust for rounding errors. During the first year that SFAS 132R was effective, firms disclose historical prior-year asset allocations in addition to the current-year allocations. I hand-collect the historical allocations from the financial statements in the first year that SFAS 132R is effective.

The asset allocations used in my study represent actual asset allocations as opposed to long-term target allocations. As a practical matter, I cannot use target allocations, which are not available for the year before SFAS 132R because SFAS 132R did not mandate target allocation disclosures for a past period. It would be conceptually appealing to use target asset allocations because they represent long-term average asset allocations and theoretically might be more closely aligned (relative to actual allocations) with the ERR. Actual allocations are specific to a particular point in time (i.e., the end of the fiscal period) and, therefore, can be affected by realized returns during the period, so they might not map well into the ERR. However, there are shortcomings associated with target allocations. Because they represent "targets," firms can have large deviations between their target allocations and actual allocations without showing any evidence of long-term convergence. Furthermore, target allocations are not verifiable, whereas actual allocations are. For my study, I use actual asset allocations.

To construct a proxy for the alignment between the ERR and asset allocation, I first identify firms that are likely using inflated ERRs to manage earnings in the year prior to SFAS 132R. I assume that ERRs are explained by pension asset composition (SFAS 87, ¶45). I regress the ERR on the proportions of the firms' pension assets invested in equities, bonds, real estate, and other unspecified securities.[12] Specifically, I regress ERR in year $t+1$ on asset allocation percentages in year t, because firms choose the ERR at the beginning of the fiscal year based on asset allocation at the end of the prior fiscal year. An alternative but equivalent perspective for the latter statement is that firms choose ERRs at the end of the previous fiscal year based on asset allocation at the end of

[11] The surveys are sent out during the summer, and pension sponsors are instructed to provide asset allocation information as of the most recent fiscal year-end date. After receiving the survey responses, P&I compiles and distributes its results every January. Thus, for a December 31 firm, the asset composition at December of year t is published in the P&I report in January of year $t+2$. Per conversations with sales representatives at P&I, the average response rate is 80 percent.

[12] The specification of this regression is subject to two caveats. First, firms can lock in a high ERR by immunizing the pension plan with bonds of similar maturity. To the extent that some firms in my sample did this, the relation between percentage of assets in bonds and ERR would be weaker. Second, the four asset categories are likely measured with error, because each category is quite broad. For instance the bonds category might include both municipal bonds and junk bonds. Therefore, it is possible that assets with different risk characteristics are classified within the same asset category.

the previous fiscal year.[13] Thus, the ERR in year t+1 is based on asset allocation at the end of year t. I also control for the effect of actual returns to pension assets, measured as the actual dollar return scaled by beginning plan assets (ARR), which can also influence the choice of ERR. Because the choice of ERR is likely affected by multiple prior years' actual returns, I include the average ARR for the three years ending in year t (*ARR_AVG3YEAR*).[14] Last, I include industry indicator variables because some firms state that they use peer data when establishing the ERR. I follow the industry classification based on Biddle and Seow (1991) and Mittelstaedt et al. (1995):

$$\begin{aligned} ERR_{it+1} = {} & a1\%EQUITIES_{it} + a2\%BONDS_{it} + a3\%REAL\_ESTATE_{it} + a4\%OTHER_{it} \\ & + a5ARR\_AVG3YEAR_{it} + \sum INDUSTRY_{it} + u_{it}. \end{aligned} \tag{1}$$

Because SFAS 132R reduces the opportunity to report biased ERRs, I use all firm-years after SFAS 132R as an estimation period to obtain parameter estimates that describe the relation between ERR and asset composition in the absence of earnings management. Then, using the parameters from my estimation period, I compute prediction errors for each firm as a measure of the portion of the firm's pre-SFAS 132R ERR that is not explained by the pension asset allocation (hereafter, the unexplained ERR: *UNEXPL_ERR*).[15] More positive prediction errors suggest that the firm is using an ERR higher than the ERRs of other firms with similar pension asset compositions in the year immediately before SFAS 132R. To validate the proxy for the unexplained ERR, I manually check the values of the unexplained ERR for the six companies targeted by the SEC for aggressive pension assumptions (discussed in Section III). I find that the values of the unexplained ERR for these six companies are among the highest values in the cross-section for 2002 and 2003.

Next, I model the firm's response to SFAS 132R as a joint decision to allocate more pension assets to risky investments and/or to reduce the reported ERR to correct the pre-SFAS 132R upward bias in ERR.[16] I use equity holdings to proxy for risky investments, because equities historically have the highest mean and variance in returns among the four categories in Equation (1). Specifically, I measure the change in equity allocation (Δ*%EQUITIES*) and the change in ERR (Δ*ERR*) around SFAS 132R. The timeline below illustrates how the issuance of SFAS 132R overlaps with the chronology of when asset allocation and the ERR are disclosed by a typical December 31 year-end firm:

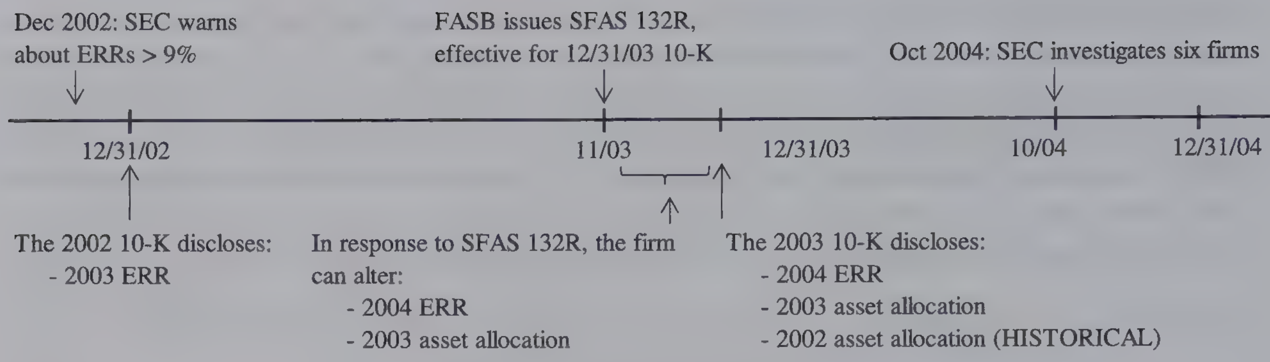
---

[13] The following example illustrates how firms choose the ERR at the beginning of the fiscal year (or equivalently, at the end of the prior fiscal year) based on asset allocation at the end of the prior fiscal year. In its 2002 annual report, Consolidated Edison states: "In addition, the Company has lowered its expected annual asset return for the plans for 2003 to 8.8 percent" (Consolidated Edison 2002, 27). Thus, at the end of 2002, the company chooses its 2003 ERR based on asset allocation at the end of 2002.

[14] Instead of the three-year average ARR, I could use one-year ARR. Untabulated results for the yearly regressions of Equation (1) show that the coefficient on one-year *ARR* is negative for several years, likely due to the higher variability in one-year *ARR*.

[15] Alternatively, I can compute in-sample residuals for the year prior to SFAS 132R. However, as discussed later in Table 2, the R^2 and parameter estimates in 2002 are different from the pattern of those in all other years. One likely explanation is that the ERR and asset composition are misaligned because of instances of unrealistic ERRs assumed in 2002. Thus, using 2003 to 2008 as an out-of-sample estimation period more likely results in parameter estimates that more effectively measure the relation between the ERR and asset composition when earnings management is constrained. Table 2, Panel B provides the computation.

[16] I predict cross-sectional variation in firms' responses and stronger responses are likely to occur for firms with more extreme misalignment between the ERR and asset composition. As an example of extreme misalignment, Midwest Air Group reported in its 2003 10-K an ERR of 9 percent in both 2002 and 2003. However, its asset compositions were drastically different across years. In 2002, Midwest invested in 100 percent bonds. Then, in 2003 (the first year of SFAS 132R), Midwest reported 69 percent equities and 31 percent bonds. Recall that firms were required to retroactively disclose their 2002 asset compositions and were unable to alter them, but they were able to alter their 2003 asset allocations in response to SFAS 132R. I expect stronger firm responses to be in cases such as Midwest Air.



Based on this timeline, I measure *Δ%EQUITIES* around SFAS 132R as the difference between the 2002 and 2003 equity allocations; firms can alter asset allocation up to the end of the fiscal year. In contrast, I measure *ΔERR* around SFAS 132R as the difference between the 2003 and 2004 ERRs because firms choose the ERR at the beginning of the fiscal year, or equivalently, at the end of the prior fiscal year.[17] Also denoted on the timeline are the two SEC actions discussed earlier. In December 2002, the SEC publicly warned that it might challenge ERRs above 9 percent. In October 2004, the SEC launched an investigation into the pension assumptions of six firms. As illustrated by the timeline, the October 2004 event is unlikely to explain the changes in asset allocation and ERR around SFAS 132R. However, it is possible that firms' delayed response to the December 2002 event might explain the observed changes. I address this alternative explanation in Section V.

I examine both changes in asset allocation and ERR in response to SFAS 132R. In Section V, descriptive statistics show that changes in asset allocation are more common than changes in ERR. Nonetheless, I examine both changes because they are likely jointly determined by a firm. Due to the joint nature of the decision, I use two-stage least squares.[18] In the first stage, I separately regress *Δ%EQUITIES* and *ΔERR* on a set of instruments.[19] I obtain the fitted values *ΔERR_HAT* and *Δ%EQUITIES_HAT* from the first-stage regressions. In the second stage, I estimate Equations (2) and (3) below using *ΔERR_HAT* and *Δ%EQUITIES_HAT*, respectively, as explanatory variables along with other control variables:

$$\Delta\%EQUITIES_i = c0 + c1UNEXPL\_ERR_i + c2\Delta ERR\_HAT_i + \sum Controls_i + u_i; \tag{2}$$

$$\Delta ERR_i = d0 + d1UNEXPL\_ERR_i + d2\Delta\%EQUITIES\_HAT_i + \sum Controls_i + u_i. \tag{3}$$

---

[17] For a 12/31 firm, 12/31/03 is the first fiscal year-ending after 12/15/03 (the effective date of SFAS 132R). My sample includes non-12/31 firms. For firms with fiscal year-end dates between 12/15 and 12/30, fiscal 2003 is the first fiscal year when SFAS 132R is effective, so the measurement of *Δ%EQUITIES* and *ΔERR* is identical to 12/31 firms. However, for firms with fiscal year-end dates between 1/1 and 12/14, fiscal 2004 is the first fiscal year when SFAS 132R is effective. Thus, the measurement of *Δ%EQUITIES* is measured as the change between 2003 and 2004, and *ΔERR* is measured as the change between 2004 and 2005. I further discuss the treatment of non-12/31 firms in Section V. Approximately 80 percent of my sample consists of 12/31 firms.

[18] If the ERR and asset allocation are endogenous, then only two-stage least squares estimates are consistent and efficient (Kennedy 1992).

[19] I set up *W*, a matrix of instruments that includes all the exogenous variables in the system of simultaneous equations. That is, *W* includes the union of all unique explanatory variables (the variables are described later) from the second-stage Equations (2) and (3), except the fitted values. Specifically, *W* includes *UNEXPL_ERR*, *ARR*, *ARR_AVG3YEAR*, *FUNDING*, *HORIZON*, *OPER_RISK*, *TAX*, *EQUITY_ RET*, *BOND_ RET*, *FIRM_SIZE*, *DB_SIZE*, *ΔINTEREST_RATE*, *EQUITY_RET_EXP*, and *ERR_LESS_PEER_AVG*. In the first stage, I regress (i) *Δ%EQUITIES* on *W* to obtain *Δ%EQUITIES_HAT* and (ii) *ΔERR* on *W* to obtain *ΔERR_HAT*. I assume that *W* is exogenous and satisfies the predetermined condition. The fitted values are purged of endogeneity. The timing of each of the variables in the *W* matrix is discussed later.

H1 predicts that firms that use inflated ERRs prior to SFAS 132R are more likely to increase their asset allocation toward equity securities after SFAS 132R ($c1 > 0$ in Equation (2)). H2 predicts that firms that use inflated ERRs prior to SFAS 132R are more likely to reduce ERRs after SFAS 132R ($d1 < 0$ in Equation (3)). I expect $c2 > 0$ and $d2 > 0$ in the equations above if firms use reductions in ERR and increases in risky asset allocation as substitute mechanisms to correct the pre-SFAS 132R upward bias in ERRs.

Equation (2) controls for several other determinants of pension asset allocation. First, actual returns to the pension assets during the fiscal year can influence the ending balance of a firm's asset allocation. I control for the contemporaneous actual returns to pension assets *ARR*, measured as the actual dollar return scaled by beginning plan assets. Second, Harrison and Sharpe (1983) analytically show that firms should either (1) overfund the pension and invest in bonds to maximize the tax benefits of holding bonds or (2) underfund the pension and invest in equities to maximize the value of the put option provided by the Pension Benefit Guaranty Corporation (PBGC) on unfunded obligations. Thus, the funding level can impact the optimal asset allocation. I measure prior-year *FUNDING* as (Plan Assets − PBO)/MVE. Third, to match pension assets and obligations, plans with young (old) employees can invest more in stocks (bonds) because stocks are more correlated with salary increases than bonds.[20] Following Amir et al. (2010), I measure prior-year *HORIZON* as log (PBO/service cost). Fourth, firms tend to offset high operating risk by investing more in bonds (Friedman 1983; Bodie et al. 1987). I measure *OPER_RISK* as the book value-deflated standard deviation of operating earnings from the past five years, ending in the prior year. Fifth, firms subject to higher tax rates have greater incentives to allocate pension assets to bonds, which are more heavily taxed (Amir et al. 2010). Following Graham (1996), I measure prior-year *TAX* as the simulated marginal tax rate before interest expense obtained from John Graham's website.[21] Sixth, if a firm does not actively rebalance its pension asset mix and the returns to its high-risk assets outperform returns to its low-risk assets, then the relative ending balances of asset values will appear as if the firm shifted assets toward high-risk securities, and *vice versa*. I include contemporaneous stock (bond) market returns as proxies for the performance of the firm's high-risk (low-risk) pension assets. I measure *EQUITY_RET* as annual returns to the CRSP value-weighted market index and *BOND_RET* as returns to the Vanguard Total Bond Market Index, both measured during the fiscal year.

Next, I control for the following determinants of ERR in Equation (3). First, it is possible that firms alter their ERRs in response to prior actual returns to pension assets. Thus, I control for average ARR computed for three years ending in the most recent year observed before the ERR is chosen (*ARR_AVG3YEAR*). While one-year ARR affects the year-end asset allocation, the choice of ERR is likely determined by multiple prior years' ARR. Second, the SEC warned in 2002 that it might challenge rate-of-return assumptions above 9 percent. Due to its limited resources, the SEC is more likely to investigate large, publicly visible firms and firms with significant DB plans. To avoid

---

[20] The idea of allocating pension assets according to the age of the workforce is an alternative to the all-bonds optimal allocation strategies under Black (1980), Tepper (1981), and Merton (2006), who do not take the age of the workforce into account.

[21] The marginal tax rate is the tax rate that applies to the last dollar of the tax base and is often applied to the change in a firm's tax obligation as income rises. The simulated tax rates account for many important features of the tax code including uncertainty about taxable income, deferred taxes, the progressivity of the statutory tax schedule, net operating loss carryforwards and carrybacks, certain tax credits, and the alternative minimum tax. The simulated tax rates are based on consolidated financial statement data and use the assumption that worldwide net income is taxed at the top U.S. statutory rate. Therefore, the simulated marginal tax rates will be measured with error for firms that face a different tax rate on foreign income. These errors could influence the results observed in my study, because the pensions in my study primarily pertain to a firm's U.S. operations.

SEC scrutiny, these types of firms are more likely to reduce ERRs, even if their ERRs might not be unreasonable with respect to asset composition. I measure prior-year firm size (*FIRM_SIZE*) as the log of market value of equity and prior-year pension size (*DB_SIZE*) as the log of PBO. Third, firms likely incorporate prevailing interest rates into their estimates of ERR, where increases in interest rates lead to increases in expected returns. I obtain monthly AAA-rated corporate bond interest rates from the Federal Reserve, and measure the contemporaneous change in interest rates during the fiscal year (*ΔINTEREST_RATE*). Fourth, under SFAS 87, the ERR represents the long-term expected (future) rate of return, such that ERRs are higher when managers are optimistic about the future performance of their investments. I proxy for expected future returns using realized future returns, under the assumption of perfect foresight. I measure future realized returns as the three-year returns to the CRSP value-weighted market index in the three years beginning after the current fiscal year (*EQUITY_RET_EXP*).[22] Fifth, to avoid public scrutiny, a firm with an ERR above those of its peer firms can reduce the ERR. I measure *ERR_LESS_PEER_AVG* as the firm's prior-year ERR minus the prior-year industry mean ERR.

For clarity, I briefly summarize the timing of each variable. The dependent variables in Equations (2) and (3) are changes measured as the firm's choice made at the end of 2003 minus the firm's choice made at the end of 2002 (i.e., 2003 *%EQUITIES* minus 2002 *%EQUITIES*, and 2004 *ERR* minus 2003 *ERR*). I measure *FUNDING*, *HORIZON*, *TAX*, *OPER_RISK*, *FIRM_SIZE*, *DB_SIZE*, and *ERR_LESS_PEER_AVG* at the end of 2002 because the level of the variable at the end of 2002 likely influences changes in firm behavior from 2002 to 2003. I measure *ARR*, *EQUITY_RET*, and *BOND_RET* during 2003 because they can each mechanically influence the 2003 end-of-year asset allocation. I measure *ΔINTEREST_RATE* as the change in interest rates during 2003, which can influence the ERR chosen at the end of 2003. I measure *EQUITY_RET_EXP* over the period 2004, 2005, and 2006 (i.e., three years beginning after the firm's decisions in 2003 are measured). I measure *ARR_AVG3YEAR* as the average ARR during 2001, 2002, and 2003 such that the last year included in the average is the most recent year observed before the firm's decision at the end of 2003 is made.

V. RESULTS

Descriptive Statistics

Table 1, Panel A presents descriptive statistics for my sample during 2002 to 2008. Except for the asset allocation percentages, I winsorize all variables to the first and 99th percentiles to reduce the effect of outliers. The top four rows represent the percentage of pension assets invested in various securities. Many firms invest predominantly in equities and fixed-income securities (medians of 61 percent and 33 percent, respectively). The mean and median ERRs during my sample period are 7.89 percent and 8.1 percent, respectively. The set of firms with DB pensions includes many of the largest firms in the population of all traded firms, with a median (mean) market capitalization of $1.5b ($6.6b). Pension assets and liabilities are economically nontrivial (medians of $125m and $167m, respectively). Aggregate mean equity returns (*EQUITY_ RET*) are lower than aggregate bond returns (*BOND_ RET*) during the sample period, with equity returns having higher standard deviations than bond returns—consistent with the notion that equity

[22] The values of *EQUITY_RET*, *BOND_RET*, *ΔINTEREST_RATE*, and *EQUITY_RET_EXP* are the same for all firms with the same fiscal year-ends, because the first three variables are measured during the fiscal year and the fourth is measured beginning at the end of the fiscal year. For instance, all firms with 12/31 year ends will have the same values for each of the variables.

TABLE 1

Descriptive Statistics

Panel A: All Years Aggregated (2002 to 2008)

Variable	P1	P5	P25	P50	P75	P95	P99	Mean	Std. Dev.
%EQUITIES	0.00%	23.00%	52.00%	61.00%	69.00%	78.70%	96.00%	58.20%	17.00%
%BONDS	0.00%	14.00%	26.30%	33.00%	40.70%	65.00%	100.00%	35.10%	16.20%
%REAL_ESTATE	0.00%	0.00%	0.00%	0.00%	0.00%	9.00%	14.50%	1.40%	3.40%
%OTHER	0.00%	0.00%	0.00%	1.00%	6.00%	24.00%	57.80%	5.30%	11.30%
ERR	2.70%	5.50%	7.50%	8.10%	8.50%	9.00%	9.60%	7.89%	1.19%
PENSION ASSETS/MVE	0.00	0.01	0.04	0.10	0.26	1.22	12.03	0.41	1.39
MVE ($m)	5.39	41	339	1,502	5,736	39,823	53,846	6,632	12,419
UNSCALED PENSION ASSETS ($m)	0.83	3	26	125	594	5,799	20,114	1,087	2,967
UNSCALED PBO ($m)	1.15	5	37	167	717	7,355	22,906	1,305	3,478
UNSCALED ABO ($m)	1.04	5	38	167	700	6,493	20,400	1,179	3,076
FUNDING	−4.24	−0.38	−0.06	−0.02	0.00	0.02	0.14	−0.12	0.49
HORIZON	1.85	2.47	3.12	3.54	3.98	5.15	6.57	3.62	0.81
TAX	0.00	0.05	0.34	0.35	0.35	0.36	0.38	0.30	0.10
OPER_RISK	0.01	0.02	0.03	0.06	0.11	0.35	1.28	0.11	0.18
EQUITY_RET	−0.39	−0.38	−0.01	0.08	0.16	0.33	0.39	0.05	0.21
BOND_RET	0.00	0.02	0.03	0.04	0.06	0.08	0.10	0.04	0.02
ARR	−0.32	−0.25	0.02	0.08	0.12	0.21	0.28	0.05	0.13
ARR_AVG3YEAR	−0.09	−0.05	0.01	0.07	0.10	0.14	0.19	0.06	0.06
INTEREST_RATE	5.05%	5.05%	5.32%	5.47%	5.62%	6.21%	6.21%	5.47%	0.28%
ΔINTEREST_RATE	−1.05%	−0.59%	−0.44%	−0.15%	−0.05%	0.17%	0.93%	−0.19%	0.32%
EQUITY_RET_EXP	−0.39	−0.24	−0.14	0.39	0.46	0.52	0.61	0.16	0.32
ERR_LESS_PEER_AVG	−4.25%	−1.88%	0.04%	0.74%	1.28%	2.02%	3.86%	0.54%	1.25%

%EQUITIES, *%BONDS*, *%REAL_ESTATE*, and *%OTHER* are the percentage of pension plan assets invested in equity securities, fixed-income securities (including cash), real estate, and all other unspecified securities, respectively. *ERR* is the expected rate of return on pension plan assets. (I use the higher of Compustat data item PPROR or PPRORMAX, because PPROR is the simple average assumed rate of return for both domestic and foreign pension plans, whereas PPRORMAX is the highest assumed rate of return. Generally, PPRORMAX applies to the U.S. rate of return).

Variable Definitions:
PENSION ASSETS/MVE = pension assets scaled by market capitalization;
MVE = market capitalization;
UNSCALED_PENSION_ASSETS = unscaled pension plan assets (sum of Compustat data items PPLAO and PPLAU);
UNSCALED_PBO = unscaled projected benefit obligation (sum of Compustat data items PBPRO and PBPRU);
UNSCALED_ABO = unscaled accumulated benefit obligation (sum of Compustat data items PBACO and PBACU);
FUNDING = *UNSCALED_PENSION_ASSETS* minus *UNSCALED_PBO* scaled by market capitalization;
HORIZON = investment horizon for pension assets measured as log(PBO/Service Cost);
TAX = marginal tax rate before interest expense;
OPER_RISK = standard deviation of the past five years of operating income, scaled by the book value of equity;
EQUITY_RET = annual return on the CRSP value-weighted index, measured during the fiscal year;
BOND_RET = annual return on the Vanguard Total Bond Index, measured during the fiscal year;
ARR = current-year actual dollar return on pension assets scaled by beginning pension assets;
ARR_AVG3YEAR = three-year average of *ARR*, ending in the current year;
INTEREST_RATE = monthly interest rate for Moody's AAA-rated corporate bonds obtained from the Federal Reserve;
ΔINTEREST_RATE = change in the interest rate from the beginning of the fiscal year to the end of the fiscal year;
EQUITY_RET_EXP = proxy for expected future return, and is measured as the three-year annual return on the CRSP value-weighted index, measured during the three years after the fiscal year-end date; and
ERR_LESS_PEER_AVG = firm's ERR minus the median ERR for the same industry, where industry is based on Biddle and Seow (1991) and Mittelstaedt et al. (1995).

(continued on next page)

TABLE 1 (continued)

Panel B: By Year

Row #		2002	2003	2004	2005	2006	2007	2008
1	Number of observations	877	1,176	1,178	1,294	1,256	1,233	1,222
2	*% EQUITIES*							
	Mean	58.2%	**61.4%**	61.5%	62.0%	61.8%	59.0%	50.8%
	Median	60.0%	**63.4%**	64.0%	64.7%	64.0%	61.5%	53.6%
	Std. Dev.	14.7%	**15.3%**	15.3%	15.0%	14.7%	15.5%	16.8%
3	*% BONDS*							
	Mean	34.2%	**32.1%**	32.7%	32.5%	32.2%	34.4%	41.1%
	Median	34.0%	**31.0%**	31.0%	31.0%	30.7%	32.7%	39.2%
	Std. Dev.	13.4%	**14.1%**	14.3%	14.2%	13.9%	14.9%	16.9%
4	*% REAL ESTATE*							
	Mean	1.1%	**1.0%**	1.1%	1.3%	1.5%	1.6%	1.7%
	Median	0.0%	**0.0%**	0.0%	0.0%	0.0%	0.0%	0.0%
	Std. Dev.	2.7%	**2.6%**	2.6%	3.1%	3.6%	3.7%	3.9%
5	*% OTHER*							
	Mean	6.6%	**5.4%**	4.7%	4.2%	4.5%	5.0%	6.4%
	Median	2.0%	**1.0%**	1.0%	0.8%	0.3%	0.4%	0.5%
	Std. Dev.	12.3%	**11.6%**	10.1%	8.7%	9.5%	10.4%	12.2%
6	*ERR*							
	Mean	8.58%	8.19%	**8.00%**	7.93%	7.86%	7.82%	7.74%
	Median	8.75%	8.50%	**8.20%**	8.25%	8.00%	8.00%	8.00%
	Std. Dev.	0.74%	1.08%	**1.17%**	1.20%	1.17%	1.13%	1.16%
7	% Firms with $\Delta ERR > 0$	3%	4%	**7%**	6%	9%	12%	10%
	% Firms with $\Delta ERR < 0$	52%	38%	**27%**	26%	28%	17%	24%
	% Firms with $\Delta ERR = 0$	45%	58%	**66%**	68%	64%	71%	66%
8	% Firms with $\Delta EQUITIES > 0$	NA	**76%**	50%	48%	45%	25%	14%
	% Firms with $\Delta EQUITIES < 0$	NA	**19%**	38%	38%	40%	64%	81%
	% Firms with $\Delta EQUITIES = 0$	NA	**5%**	12%	13%	15%	11%	5%
9	*ΔERR*							
	Mean	−0.36%	−0.21%	**−0.10%**	−0.09%	−0.07%	−0.02%	−0.07%
	Median	−0.10%	0.00%	**0.00%**	0.00%	0.00%	0.00%	0.00%
	Std. Dev.	0.51%	0.46%	**0.41%**	0.40%	0.50%	0.45%	0.43%
10	*Δ%EQUITIES*							
	Mean	NA	**4.77%**	0.79%	0.33%	−0.03%	−2.93%	−8.05%
	Median	NA	**4.00%**	0.10%	0.00%	0.00%	−2.00%	−7.00%
	Std. Dev.	NA	**9.92%**	9.58%	7.85%	7.80%	9.07%	10.42%
11	*ARR_AVG3YEAR*							
	Mean	−3.3%	**1.7%**	6.9%	12.2%	10.1%	8.4%	−0.9%
	Median	−3.8%	**1.7%**	6.7%	12.5%	10.2%	8.7%	−1.3%
	Std. Dev.	4.1%	**3.8%**	3.2%	3.2%	2.7%	3.4%	3.7%
12	*ARR*							
	Mean	−6.7%	**17.4%**	10.0%	9.1%	11.1%	4.8%	−18.8%
	Median	−8.3%	**17.8%**	10.1%	8.4%	11.3%	6.4%	−22.0%
	Std. Dev.	8.4%	**6.0%**	3.8%	4.6%	4.0%	8.0%	11.2%

(continued on next page)

TABLE 1 (continued)

Row #		2002	2003	2004	2005	2006	2007	2008
13	*EQUITY_RET*							
	Mean	−13.6%	**30.2%**	12.4%	9.1%	16.4%	1.4%	−33.1%
	Median	−20.8%	**33.1%**	13.0%	7.3%	16.2%	7.3%	−38.3%
	Std. Dev.	15.4%	**7.4%**	2.2%	3.3%	2.1%	11.3%	13.2%

Asset allocations are chosen at the end of the fiscal year, whereas ERRs are chosen at the beginning of the fiscal year (or equivalently, the end of the prior fiscal year). For the variables in Rows 2–10, I bold the year in which the variable would be impacted by managerial responses to SFAS 132R. For the variables in Rows 11–13, I bold the year in which the variable can potentially impact *Δ%EQUITIES* and *ΔERR*, even in the absence of managerial responses to SFAS 132R.

securities are riskier investments than bonds. As expected, when compared to the one-year *ARR*, the three-year average ARR (*ARR_AVG3YEAR*) has a lower standard deviation and is less prone to the extreme negative actual returns that are common during my sample period.

Before reporting descriptive statistics by year, I discuss the treatment of non-12/31 firms for the purpose of performing analyses by year. As mentioned in Section IV, a complication arises in obtaining an appropriate cross-sectional aggregation of 12/31 firms and non-12/31 firms. SFAS 132R became effective for fiscal years ending after December 15, 2003. Thus, for a 12/31 firm, 2003 is the first year in which SFAS 132R is effective, and I label these firm-years as year 2003. Also, for a 12/31 firm, I label December 31, 2002 as year 2002, which represents the year immediately before SFAS 132R. Next, I find the fiscal year immediately before SFAS 132 for non-12/31 firms, and I label these observations as year 2002 along with the December 31, 2002 observations. For non-12/31 firms, the label 2002 includes fiscal year-end dates between December 15, 2002 and December 14, 2003. Collectively, these aggregated 12/31 and non-12/31 observations represent the pre-SFAS 132R firm-years. Correspondingly, the label 2003 (which represents the first year of SFAS 132R) includes fiscal year-end dates between December 15, 2003 to December 14, 2004. The label 2004 includes fiscal year-end dates between December 15, 2004 and December 14, 2005, and so on. As a result, firm-years are defined using SFAS 132R as a reference point. Specifically, firm-years are aggregated such that regardless of the fiscal year-end date, firm-years are grouped together as year 2003 if they are the first year in which SFAS 132R is effective, and grouped together as year 2004 if they are the second year in which SFAS 132R is effective, etc. For the remaining analyses, the year labels refer to the aggregation procedure above.

Table 1, Panel B reports descriptive statistics by year for select variables. For the variables in rows 2 to 10, I bold the year in which the variable is predicted to be affected by managerial responses to SFAS 132R. For the variables in rows 11 to 13, I bold the year in which the variable can potentially affect asset composition or the ERR, even in the absence of managerial responses to SFAS 132R.

Row 1 shows the sample size of firms with the required data necessary to estimate Equation (1) in each year. Rows 2 to 5 show that asset composition is fairly stable across the years. ERRs are also fairly stable over time, as shown in row 6. As discussed earlier, asset composition is chosen at the end of the year, whereas ERRs are chosen at the beginning of the year. Thus, to indicate the year of the predicted managerial response to SFAS 132R, asset composition is bolded in 2003 and ERR is bolded in 2004. Row 7 reports the proportion of firms in any given year that increase, decrease, or do not change their ERRs. Changes in ERR appear to be uncommon, and in the first year of SFAS 132R, 27 percent of firms in my sample reduce their 2004 ERRs. However, the fraction of firms that reduce their ERRs (row 7) and the mean ERR reduction (row 9) in 2002 and 2003 are larger than in

2004. The reductions in 2002 and 2003 are more likely due to SEC pressures or other factors than in response to SFAS 132R. In addition, the fraction of firms that reduce ERRs and the mean reductions in 2005 and 2006 are similar to that in 2004. Overall, it appears that the results for changes in ERR are weaker than those for changes in equity allocation, as discussed next. Row 8 reports the proportion of firms in any given year that increase, decrease, or do not change their equity allocation. In the first year of SFAS 132R, 76 percent of firms in the sample increased their asset allocation in equities. The magnitude of this proportion is large compared to the other years in the sample period. Rows 9 and 10 show that, in first year of SFAS 132R, firms on average reduced their ERRs and increased their equity allocation. Row 11 shows that the low median *ARR_AVG3YEAR* of 1.7 percent can contribute to firms' decision to reduce their ERR. Rows 12 to 13 report actual returns to pension assets and stock market index returns, respectively. The median one-year ARR of 17.8 percent and the median stock market index return of 33.1 percent in 2003 can mechanically affect the ending balance in the pension assets held in equities, even in the absence of any active portfolio rebalancing. Thus, these results show the importance of my controls for actual returns and equity index returns in multivariate tests, as discussed later.

Descriptive Analysis of SFAS 132R

Table 2, Panel A reports results of regressing ERR on asset composition using Equation (1) for each year 2002 to 2008, controlling for prior actual returns and industry. The coefficients on each asset category can be interpreted as the average implied expected return to that particular asset category. For instance, the coefficient of 8.26 for *%EQUITIES* in 2004 implies that during 2004 managers on average expected an 8.26 percent return on equity securities. The results of the t-tests on the right side of the table support the idea that equities have statistically higher implied expected returns than bonds, real estate, and other securities. Furthermore, the cross-sectional model in Equation (1) appears to have reasonable explanatory power for the ERR.[23] Overall, the results suggest that asset allocation is an important determinant of the ERR.

Several findings are noteworthy about the 2002 results. First, the R^2 in 2002 is only 17.7 percent, whereas R^2 from the other six years ranges from 20.1 percent to 26.7 percent. Thus, it appears that asset composition has less explanatory power for the ERR before asset composition was required to be disclosed under SFAS 132R. Second, 2002 is the only year in which there is a negative relation between ERR and prior actual returns, suggesting that before SFAS 132R managers likely exercise discretion in choosing ERRs that are not positively correlated with prior actual returns. Third, 2002 is the only year in which the implied expected returns to real estate—or any other category, for that matter—are higher than equities (10.07 > 8.78). This suggests that before SFAS 132R, the differences in implied expected returns to the highest-risk investment category (i.e., equities) versus those to the lower-risk categories do not explain ERRs as well as in the years after SFAS 132R.

In an untabulated analysis, I define an indicator variable *POST132R* equal to 1 for all years after SFAS 132R (2003 to 2008), and equal to 0 for 2002, and I interact *POST132R* with the asset allocation and actual return variables in Equation (1). Results of this test show that the differences in implied expected returns between equity securities and the three other categories become

---

[23] I estimate Equation (1) without an intercept because *%EQUITIES*, *%BONDS*, *%REAL_ESTATE*, and *%OTHER* add up to 100 percent, such that any one of the asset categories can be expressed as a linear combination of the other three asset categories. In any regression without an intercept, R^2 is redefined and cannot be interpreted in the traditional sense as a measure of explanatory power. Thus, to provide a measure of explanatory power, the R^2 reported in this table is computed from untabulated regressions of a modified version of Equation (1) in which I drop the variable *%EQUITIES* and include an intercept.

TABLE 2

Results of Estimating Equation (1)—*ERR* on Asset Allocation

Panel A: Regressions by Year

$$ERR_{it+1} = a1\%EQUITIES_{it} + a2\%BONDS_{it} + a3\%REAL\_ESTATE_{it} + a4\%OTHER_{it} + a5ARR\_AVG3YEAR_{it} + \sum INDUSTRY \quad (1)$$

	Variable	Coefficient	t-stat	p-value	Tests of Coefficients Against *%Equities*	p-value
2002	*%EQUITIES*	8.78	89.9	<0.001		
	%BONDS	7.51	59.0	<0.001	Test: a1 = a2	<0.001
	%REAL_ESTATE	10.07	11.4	<0.001	Test: a1 = a3	0.077
	%OTHER	7.41	41.0	<0.001	Test: a1 = a4	<0.001
	ARR_AVG3YEAR	−1.50	−2.4	0.016		
	n = 877					
	R^2 = 17.7%					
2003	*%EQUITIES*	8.86	95.2	<0.001		
	%BONDS	6.95	45.1	<0.001	Test: a1 = a2	<0.001
	%REAL_ESTATE	8.23	7.4	<0.001	Test: a1 = a3	0.139
	%OTHER	4.47	18.5	<0.001	Test: a1 = a4	<0.001
	ARR_AVG3YEAR	2.25	2.9	0.004		
	n = 1,176					
	R^2 = 24.5%					
2004	*%EQUITIES*	8.26	66.6	<0.001		
	%BONDS	5.80	35.5	<0.001	Test: a1 = a2	<0.001
	%REAL_ESTATE	5.61	4.8	<0.001	Test: a1 = a3	0.013
	%OTHER	4.78	16.0	<0.001	Test: a1 = a4	<0.001
	ARR_AVG3YEAR	8.90	8.7	<0.001		
	n = 1,178					
	R^2 = 24.3%					
2005	*%EQUITIES*	8.23	44.4	<0.001		
	%BONDS	6.05	37.6	<0.001	Test: a1 = a2	<0.001
	%REAL_ESTATE	6.39	6.4	<0.001	Test: a1 = a3	0.033
	%OTHER	5.26	15.0	<0.001	Test: a1 = a4	<0.001
	ARR_AVG3YEAR	3.08	2.9	0.004		
	n = 1,294					
	R^2 = 20.1%					
2006	*%EQUITIES*	8.08	49.5	<0.001		
	%BONDS	5.40	34.4	<0.001	Test: a1 = a2	<0.001
	%REAL_ESTATE	5.18	6.5	<0.001	Test: a1 = a3	<0.001
	%OTHER	5.43	18.0	<0.001	Test: a1 = a4	<0.001
	ARR_AVG3YEAR	6.99	6.1	<0.001		
	n = 1,256					
	R^2 = 25.0%					
2007	*%EQUITIES*	8.52	62.5	<0.001		
	%BONDS	5.39	36.1	<0.001	Test: a1 = a2	<0.001
	%REAL_ESTATE	6.21	7.4	<0.001	Test: a1 = a3	0.003
	%OTHER	5.46	18.8	<0.001	Test: a1 = a4	<0.001
	ARR_AVG3YEAR	4.93	5.2	<0.001		
	n =1,233					
	R^2 = 26.7%					

(continued on next page)

TABLE 2 (continued)

	Variable	Coefficient	t-stat	p-value	Tests of Coefficients Against *%Equities*	p-value
2008	*%EQUITIES*	9.30	78.0	<0.001		
	%BONDS	5.65	44.4	<0.001	Test: a1 = a2	<0.001
	%REAL_ESTATE	7.00	8.4	<0.001	Test: a1 = a3	0.004
	%OTHER	5.72	22.6	<0.001	Test: a1 = a4	<0.001
	ARR_AVG3YEAR	5.70	6.1	<0.001		
	n = 1,222					
	$R^2 = 26.1\%$					

All variables are defined in Table 1.
SIC indicator variables are included but not tabulated. Industry classification is based on Biddle and Seow (1991) and Mittelstaedt et al. (1995), and is described in Table 3, Panel B. The following example illustrates the timing of the variables above. For the year labeled as 2003, the dependent variable is the 2004 ERR (chosen by the firm at the end of 2003). The asset allocation variables are measured at the end of 2003. *ARR_AVG3YEAR* is the average ARR during 2001, 2002, and 2003, because the average ARR in years ending in 2003 likely influences the choice of ERR made at the end of 2003.

Panel B: Pooled Regression Using 2003–2008 as Estimation Period

$$ERR_{it+1} = a1\%EQUITIES_{it} + a2\%BONDS_{it} + a3\%REAL\_ESTATE_{it} + a4\%OTHER_{it} + a5ARR\_AVG3YEAR_{it} + \sum INDUSTRY \quad (1)$$

Coefficient	Variable	Coefficient	t-stat	p-value
a1	*%EQUITIES*	8.859	92.880	<0.001
a2	*%BONDS*	6.055	42.310	<0.001
a3	*%REAL_ESTATE*	6.977	9.250	<0.001
a4	*%OTHER*	5.407	13.060	<0.001
a5	*ARR_AVG3YEAR*	0.631	2.360	0.019
n = 7,359				
$R^2 = 23.5\%$				

The above results are from a pooled regression using all firm-years in 2003–2008 as an estimation period. These parameters are used to compute the variable *UNEXPL_ERR*, which represents prediction errors in 2002 (the year prior to SFAS 132R) calculated as follows using asset allocation in 2002 and three-year average ARR ending in 2002, where the coefficients denoted with hats are the parameter estimates using the estimation period as reported above: Pre-SFAS 132R ERR minus ($\hat{a}1 \times \%EQUITIES + \hat{a}2 \times \%BONDS + \hat{a}3 \times \%REAL\_ESTATE + \hat{a}4 \times \%OTHER + \hat{a}5 \times ARR\_AVG3YEAR$). For example, suppose a firm reports a pre-SFAS 132R ERR of 9 percent, and has a prior three-year average ARR of 10 percent. The firm also reports 60 percent, 20 percent, 15 percent, and 5 percent allocated to Equities, Bonds, Real Estate, and Other, respectively, for 2002. Therefore, *UNEXPL_ERR* for this firm has a value of $9 - [(8.859 \times 60\%) + (6.055 \times 20\%) + (6.977 \times 15\%) + (5.407 \times 5\%) + (0.631 \times 10\%)] = 1.094$. That is, under the model specified in Equation (1), this firm's pre-SFAS 132R ERR of 9 percent is 1.094 percent higher than justified by its asset allocation and prior actual returns, controlling for industry. Standard errors are clustered by firm. SIC indicator variables are included but not tabulated.

statistically larger after SFAS 132R. Overall, the results suggest that after SFAS 132R, there were significant changes in the relation between the ERR and asset composition.[24]

Table 2, Panel B reports the results of estimating Equation (1) using the pooled sample in 2003 to 2008 as a post-SFAS 132R estimation period. The parameters from Table 2, Panel B are used to compute prediction errors for 2002: the portion of the ERR not explained by asset allocation and prior actual returns (*UNEXPL_ERR*).

The formal tests of H1 and H2 require additional variables, and the data requirements reduce the sample size. Table 3, Panel A tabulates the sequential reduction in sample size as a result of each data requirement. Row 1 reports the number of firms with data to estimate Equation (1) in 2003. Of the 1,176 firms in 2003, there are 887 firms with the requisite data to estimate Equation (1) in 2002, as reported in row 2.[25] The remaining rows are presented in increasing order of data restrictiveness. Rows 3 to 13 show that no observations are deleted for a majority of the variables. The two largest losses in data are for the computation of *OPER_RISK* in row 17 ($828 - 768 = 60$ firms deleted) and *TAX* in row 18 ($768 - 578 = 190$ firms deleted). To maintain sample size for my formal tests of H1 and H2, I report results without the *TAX* variable and use the sample of 768 firms in row 17 as my main sample.[26]

Table 3, Panel B reports the industry composition of three samples of interest: (1) the full initial sample of 887 firms from row 2 in Table 3, Panel A, (2) the main sample of 768 firms from row 17, and (3) the most restricted final sample of 578 firms in row 18 after requiring all data items. There is minimal industry clustering, with only about a handful of industries representing more than 5 percent of each of the three subsamples. Also, the data requirements do not appear to distort the industry composition of the sample.

Table 4 reports the univariate correlations among the variables of interest for the sample of 768 firms from row 17 of Table 3, Panel A. As preliminary support for H1 and H2, the correlation between *UNEXPL_ERR* and *Δ%EQUITIES* (*ΔERR*) is significantly positive (negative). For brevity, I discuss only the Spearman correlations for the control variables with predicted signs. *Δ%EQUITIES* is significantly correlated with *ARR*, *EQUITY_RET*, and *BOND_RET* in the predicted directions. *ΔERR* is significantly correlated with *FIRM_SIZE*, *DB_SIZE*, and *ERR_LESS_PEER_AVG* in the predicted directions.

---

[24] I perform two untabulated sensitivity tests using interaction terms with the *POST132R* indicator variable. First, I re-estimate the equation using a constant sample of firms for which data are available for all years 2002 to 2008. The coefficients and significance levels remain similar. This specification mitigates concerns that results are driven by changes in Compustat coverage, survivorship bias, or self-selection bias. Second, to test whether results for any one single year tend to be systematically different from pooled results for other years, I re-estimate the equation six times after separately redefining the *POST132R* indicator to equal 1 for all years except 2003, 2004,. . .,2008, as opposed to setting the *POST132R* indicator equal to 1 for all years except 2002. Untabulated results from these six regressions show that the coefficients on the interaction terms are generally not significant in these alternative model specifications. These results lend further support that the relation between the ERR and asset allocation in 2002 is different from that in 2003–2008.

[25] The loss of firms from row 1 to row 2 is mainly due to some firms' failure to disclose their prior-year asset allocations in the first year of SFAS 132R. Untabulated tests show that there are no statistical differences between the 887 firms that disclose prior-year allocations and the remaining firms that fail to disclose for the following measures: the level of ERR both before and after SFAS 132R, the change in ERR around SFAS 132R, and post-SFAS 132R equity allocation.

[26] Untabulated results show that *UNEXPL_ERR* remains significant in the predicted directions if using the sample of 578 firms in row 18 and including the *TAX* variable, which is not statistically significant. About 25 percent ((768 – 578)/768) of the firms in row 17 have missing marginal tax rates from John Graham's website. These observations are missing either because the firm is new to Compustat and therefore has insufficient historical data to run the simulation or otherwise has missing data for one or more of the key input variables in the simulation procedure (Graham and Mills 2008). Graham and Mills (2008, footnote 16) report that 36 percent of their sample firms have missing marginal tax rates obtained from the same source.

TABLE 3
Sample Selection

Panel A: Tabulation of Sample Size Restrictions for Various Variables

Row #		
1	Firms with Nonmissing Data For Equation (1) in 2003	1,176
2	Firms with Nonmissing Data For Equation (1) in 2002	887
3	Firms with Nonmissing *UNEXPL_ERR*	887
4	Firms with Nonmissing *ΔERR_HAT*	887
5	Firms with Nonmissing *ΔEQUITIES_HAT*	887
6	Firms with Nonmissing *ARR*	887
7	Firms with Nonmissing *ARR_3YEAR_AVG*	887
8	Firms with Nonmissing *EQUITY_RET*	887
9	Firms with Nonmissing *BOND_RET*	887
10	Firms with Nonmissing *DB_SIZE*	887
11	Firms with Nonmissing *ΔINTEREST_RATE*	887
12	Firms with Nonmissing *EQUITY_RET_EXP*	887
13	Firms with Nonmissing *ERR_LESS_PEER_AVG*	887
14	Firms with Nonmissing *FIRM_SIZE*	879
15	Firms with Nonmissing *FUNDING*	879
16	Firms with Nonmissing *HORIZON*	828
17	Firms with Nonmissing *OPER_RISK*	768
18	Firms with Nonmissing *TAX*	578

The formal tests of H1 and H2 require additional variables, and the data requirements reduce the sample size. This table tabulates the sequential reduction in sample size as a result of each data requirement. Row 1 reports the number of firms with data to estimate Equation (1) in 2003. Of the 1,176 firms in 2003, there are 887 firms with the requisite data to estimate Equation (1) in 2002. The loss of firms from row 1 to row 2 is mainly due to some firms' failure to disclose their prior-year asset allocations in the first year of SFAS 132R. Untabulated tests show that there are no statistical differences between the 887 firms that disclose prior-year allocations and the remaining firms that fail to disclose for the following measures: the level of ERR both before and after SFAS 132R, the change in ERR around SFAS 132R, and post-SFAS 132R equity allocation. The remaining rows above are presented in increasing order of data restrictiveness. To retain sample size, I report main results in Tables 4, 5, and 6 using the sample of 768 firms in row 17 above without requiring the *TAX* variable. Untabulated tests using the sample of 578 firms in row 18 above with the *TAX* variable show that the variable of interest for H1 and H2 (*UNEXPL_ERR*) remains significant in the predicted directions.

Panel B: Industry Composition

		Column I		Column II		Column III	
		Firms in Row 2 from Panel A		Firms in Row 17 from Panel A		Firms in Row 18 from Panel A	
Industry Group	SIC Codes	n	Percent of Sample	n	Percent of Sample	n	Percent of Sample
Mining	1000 1040 1211 1400	6	0.7%	6	0.8%	2	0.3%
Oil and Gas Production	1311 1381 1389 3533	34	3.8%	30	3.9%	14	2.4%
Construction	0900 1540 1600 2400 2421 2451 2452 2510 2520 2522	26	2.9%	20	2.6%	15	2.6%

(continued on next page)

TABLE 3 (continued)

		Column I		Column II		Column III	
		Firms in Row 2 from Panel A		Firms in Row 17 from Panel A		Firms in Row 18 from Panel A	
Industry Group	SIC Codes	n	Percent of Sample	n	Percent of Sample	n	Percent of Sample
Foods	2000 2013 2016 2020 2030 2040 2050 2052 2060 2065 2070 2080 2082 2090 2100 2111	47	5.3%	38	4.9%	29	5.0%
Textiles	2200 2211 2221 2250 2253	2	0.2%	2	0.3%	2	0.3%
Apparel	2300 2320	8	0.9%	6	0.8%	5	0.9%
Paper	2600 2621 2631 2640 2643	8	0.9%	8	1.0%	7	1.2%
Publishing	2711 2721 2731 2750 2761	39	4.4%	36	4.7%	31	5.4%
Chemicals	2800 2810 2820 2821	39	4.4%	32	4.2%	24	4.2%
Pharmaceuticals	2834	21	2.4%	16	2.1%	12	2.1%
Specialty Chemicals	2840 2842 2844 2851 2860 2870 2890 2891	48	5.4%	40	5.2%	29	5.0%
Petroleum Refining	2911	13	1.5%	10	1.3%	8	1.4%
Rubber, Plastic, and Leather	3011 3069 3079 3140	13	1.5%	10	1.3%	8	1.4%
Glass, Cement, and Ceramics	3220 3221 3241 3272 3290	6	0.7%	6	0.8%	5	0.9%
Steel	3310 3312	13	1.5%	8	1.0%	7	1.2%
Metalworks	3330 3350	15	1.7%	14	1.8%	10	1.7%
Metal Parts	3390 3411 3420 3430 3440 3444 3448 3452 3460 3490 3494	39	4.4%	38	4.9%	31	5.4%
Industrial Equipment	3510 3520 3523 3530 3531 3537 3540 3541	39	4.4%	38	4.9%	26	4.5%
Small Industrial Machinery	3550 3558 3560 3561 3562 3564 3580 3585 3590	31	3.5%	30	3.9%	26	4.5%
Electrical Machinery	3600 3610 3621 3630 3634 3640 3690	37	4.2%	32	4.2%	20	3.5%
Telecommunications Equipment	3651 3661 3663 3664 3665	13	1.5%	6	0.8%	3	0.5%

(continued on next page)

TABLE 3 (continued)

Industry Group	SIC Codes	Column I: Firms in Row 2 from Panel A — n	Column I: Percent of Sample	Column II: Firms in Row 17 from Panel A — n	Column II: Percent of Sample	Column III: Firms in Row 18 from Panel A — n	Column III: Percent of Sample
Electronic Components	3670 3674 3678	13	1.5%	12	1.6%	8	1.4%
Computers	3570 3680 3681 3682 3683 3688 3689	0	0.0%	0	0.0%	0	0.0%
Automobiles	3711 3714 3716	25	2.8%	22	2.9%	15	2.6%
Aircraft	3720 3721 3724 3728 3760	19	2.1%	18	2.3%	14	2.4%
Misc. Manu-facturing	3820 3822 3823 3825 3829 3841 3861 3910 3911 3931 3942 3944	27	3.0%	24	3.1%	20	3.5%
Commercial Transport	3730 4011 4213 4400 4411	27	3.0%	24	3.1%	19	3.3%
Air Transport	4511 4513 4700	2	0.2%	2	0.3%	2	0.3%
Telecommun-ications	4811 4833 4890	10	1.1%	8	1.0%	7	1.2%
Electric Utilities	4911	52	5.9%	46	6.0%	34	5.9%
Natural Gas	4922 4923 4924	33	3.7%	32	4.2%	24	4.2%
Other Utilities	4931 4932 4940 4953	44	5.0%	40	5.2%	31	5.4%
Wholesalers	5013 5051 5065 5070 5081 5094 5099 5110 5122 5140 5161 5172	25	2.8%	24	3.1%	20	3.5%
Department Stores	5311 5331	13	1.5%	12	1.6%	10	1.7%
Specialty Stores	5600 5621 5661 5812 5900 5944 5949	19	2.1%	13	1.7%	10	1.7%
Grocers	5411 5412 5912	23	2.6%	18	2.3%	15	2.6%
Financial Services	6120 6162 6199 6211 6281	13	1.5%	12	1.6%	10	1.7%
Investors	6411 6512 6552 6795 6798 6799	13	1.5%	9	1.2%	5	0.9%
Personal Services	7011 7200 7393 7810 7814 7948 7990 8062	19	2.1%	16	2.1%	12	2.1%
Business Services	7311 7321 7340 7360 7372 7374 7392 7394 7510 8911	13	1.5%	10	1.3%	8	1.4%
Total		887	100.0%	768	100.0%	578	100.0%

Industry classification is based on Biddle and Seow (1991) and Mittelstaedt et al. (1995).

TABLE 4

Correlation Matrix for Variables Used in Tests of Managerial Responses to SFAS 132R

Panel A: Correlation Matrix, Variables *ΔERR* to *OPER_RISK*

	ΔERR	*Δ%EQUITIES*	*UNEXPL_ERR*	*ARR (One Year)*	*ARR_AVG 3YEAR*	*FUND-ING*	*HORI-ZON*	*OPER_RISK*
ΔERR		0.0200	−0.1727***	0.0410	0.0017	0.0017	0.0069	0.0143
Δ%EQUITIES	−0.0032		0.2870***	0.0024	0.0003	0.0194	−0.0125	−0.0121
UNEXPL_ERR	−0.2552***	0.3047***		0.0601	−0.0081	−0.0198	0.0676*	0.0156
ARR (One Year)	0.0128	0.0711**	0.0311		0.2037***	−0.0279	0.0474*	−0.0163
ARR_AVG3YEAR	0.0263	−0.0446	−0.0403	0.1829***		0.0160	0.0440	0.0068
FUNDING	0.0690**	0.0413	−0.1400***	−0.0993***	0.0945***		−0.1238***	−0.1227***
HORIZON	−0.0371	−0.0398	0.0717*	0.0479*	0.0400	−0.3181***		0.0614*
OPER_RISK	−0.0101	−0.0369	0.0437	−0.0043	−0.0511	−0.1396***	0.1148***	
EQUITY_RET	−0.0193	0.1749***	0.1239***	0.2707***	−0.2104***	−0.0098	−0.0478*	0.0061
BOND_RET	0.0177	−0.0952***	−0.0872*	−0.1135***	0.0260	−0.0434	−0.0182	0.0063
FIRM_SIZE	−0.1124***	−0.0092	−0.0223	0.1540***	−0.0316	−0.0020	0.0596*	−0.3785***
DB_SIZE	−0.0874***	−0.0441	0.0959**	0.2155***	−0.0407***	−0.3439***	0.3602***	0.0511*
ΔINTEREST_RATE	0.0316	−0.1957***	−0.1659**	−0.2560***	0.2270***	−0.0103	0.0010	0.0083
EQUITY_RET_EXP	0.0165	−0.0579*	−0.0361	−0.1481***	0.0902***	−0.0226	0.0567*	−0.0174
ERR_LESS_PEER_AVG	−0.1509***	−0.0787**	−0.3573***	−0.2023***	−0.0248	−0.0629*	0.0473*	−0.0465

(continued on next page)

TABLE 4 (continued)

Panel B: Correlation Matrix, Variables *EQUITY_RET* to *ERR_LESS_PEER_AVG*

	EQUITY_ RET	*BOND_ RET*	*FIRM_ SIZE*	*DB_ SIZE*	*ΔINTEREST_ RATE*	*EQUITY_ RET_ EXP*	*ERR_ LESS_ PEER_ AVG*
ΔERR	−0.0048	−0.0142	−0.0550	−0.0356	−0.0261	0.0238	−0.1875***
Δ%EQUITIES	0.1262***	−0.0376	−0.0541	−0.0671**	−0.0505	−0.0357	−0.0767***
UNEXPL_ERR	0.1114***	−0.0449	−0.0436	0.0789**	−0.1016**	0.0629	−0.4935***
ARR (One Year)	0.2993***	−0.1048***	0.1457***	0.2106***	−0.1296***	−0.1158***	−0.2597***
ARR_AVG3YEAR	−0.1898***	−0.0282	−0.0623*	−0.0738**	0.1474***	0.0539*	−0.0634*
FUNDING	−0.0353**	−0.0171	0.2616***	−0.1076***	0.0171	0.0297	−0.0035
HORIZON	−0.0592**	−0.0101	0.0075	0.2992***	0.0320	−0.0536*	−0.0524*
OPER_RISK	−0.0001	0.0058	−0.3104***	−0.0489	0.0042	−0.0097	−0.0064
EQUITY_RET		0.0562***	0.0146	0.0490*	−0.5959***	−0.4038***	0.0657**
BOND_RET	0.1532***		0.0271*	0.0330	−0.3532***	−0.0788***	0.0945***
FIRM_SIZE	0.0236*	0.0249*		0.7009***	−0.0641***	0.0259*	0.0288
DB_SIZE	0.0352	0.0402	0.7231***		−0.0783**	0.0264	0.0351
ΔINTEREST_RATE	−0.6649***	0.0775***	−0.0533***	−0.0409		−0.1453	0.0363***
EQUITY_RET_EXP	−0.4016***	0.2151***	0.0427***	0.0417	−0.1018***		−0.2113***
ERR_LESS_PEER_AVG	0.0494*	0.0873***	−0.0143	0.0086	0.1115***	−0.1081***	

*, **, *** Significant at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.
Pearson correlation coefficients appear above the diagonal, and Spearman correlation coefficients appear below the diagonal.

Figure 1, Panel A (Panel B) plots the change in equity allocation (change in ERR) around SFAS 132R against the unexplained ERR. I form portfolios of 100 observations sorted on the unexplained ERR and plot the mean values of the variables for each portfolio.[27] As preliminary support for H1 and H2, Figure 1 shows visual evidence that the pre-SFAS 132R unexplained ERR is (1) positively related to the change in equity allocation and (2) negatively related to the change in ERR.

Main Results for Tests of H1 and H2

Table 5 reports the results of second-stage estimation of Equations (2) and (3).[28] Panel A shows results for a test of H1 using Equation (2). H1 predicts that the coefficient on *UNEXPL_ERR* > 0. Consistent with H1, the coefficient of 0.030 is significantly positive at $p < 0.001$, suggesting that firms allocate more pension assets to equities after SFAS 132R to correct the pre-SFAS 132R upward bias in ERRs. Of the control variables with predicted signs, *ARR* and *EQUITY_RET* are significant, suggesting that it is important to control for contemporaneous actual pension returns and equity index returns, which can also explain changes in asset allocation in the absence of active pension portfolio rebalancing. While *FUNDING* has no predicted sign, it is significantly positive, suggesting that firms with better funded pensions on average increase their equity allocations during this period.

Table 5, Panel B reports the results from estimating Equation (3) to test H2, which predicts that the coefficient on *UNEXPL_ERR* < 0. Consistent with H2, the coefficient is -0.216 ($p < 0.001$), suggesting that firms with upward biased pre-SFAS 132R ERRs reduce their ERRs after asset allocation disclosures are required. Of the control variables, *ARR_AVG3YEAR*, *FIRM_SIZE*, *ΔINTEREST_RATE*, *EQUITY_RET_EXP*, and *ERR_LESS_PEER_AVG* are statistically significant in the predicted directions, suggesting that prior actual returns, regulatory scrutiny, interest rates, expected future returns, and deviation from the industry average ERR can influence changes in the ERR.[29]

Results in Table 5 suggest that firms use both reductions in ERR and increases in equity allocation to correct the pre-SFAS 132R bias in ERRs. Significantly positive coefficients on the fitted values in Equations (2) and (3) would provide evidence that the two mechanisms are substitutes. However, the coefficients for *ΔERR_HAT* and *Δ%EQUITIES_HAT* are not significantly positive.

---

[27] I use the largest sample with the requisite data for this univariate analysis: the 887 firms in row 3 in Table 3, Panel A.

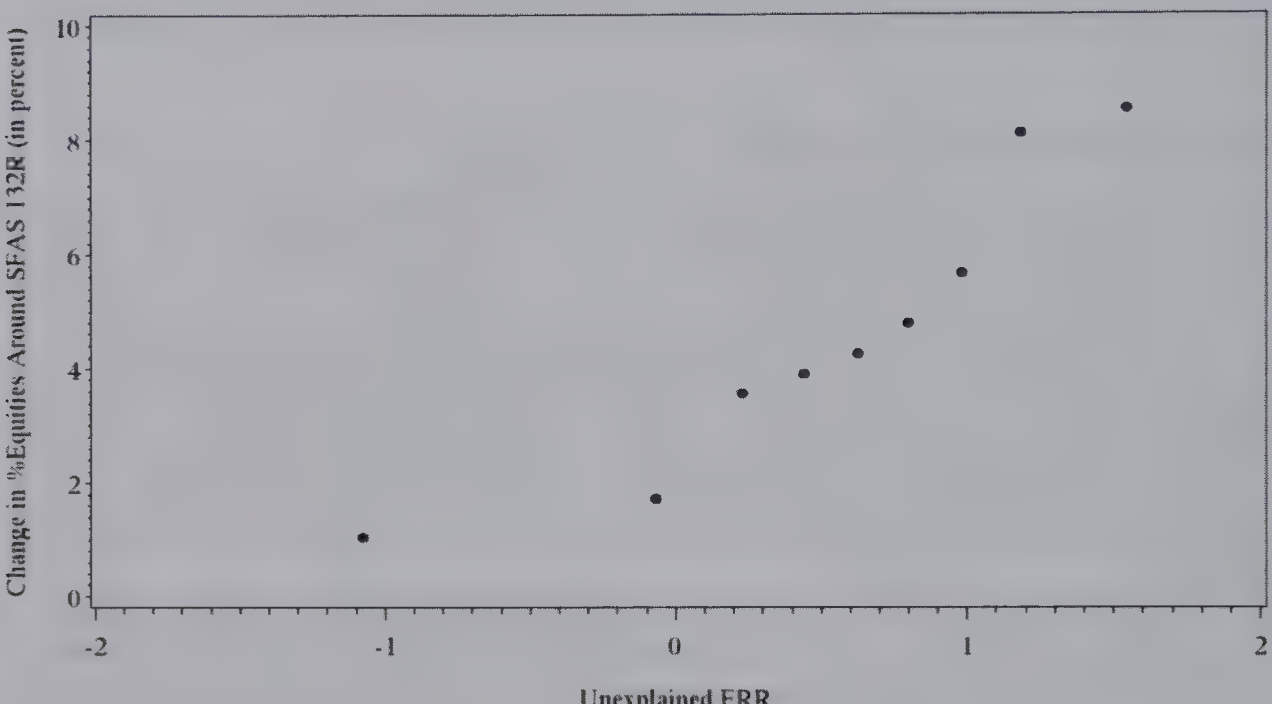
[28] In the first-stage regressions (i.e., *Δ%EQUITIES* and *ΔERR* regressed on a set of instruments), the following instruments are significantly different from zero when *Δ%EQUITIES* is the dependent variable: *UNEXPL_ERR*, *EQUITY_RET*, *DB_SIZE*, and *ERR_LESS_PEER_AVG*. The following instruments are significantly different from zero when *ΔERR* is the dependent variable: *UNEXPL_ERR*, *ARR_AVG_3YEAR*, *BOND_RET*, *FIRM_SIZE*, *DB_SIZE*, *ΔINTEREST_RATE*, and *ERR_LESS_PEER_AVG*. The first-stage R^2s are 20.0 percent and 25.1 percent, respectively. If *Δ%EQUITIES* and ΔERR are simultaneously determined, then only two-stage least squares (2SLS) estimates are consistent and efficient (Kennedy 1992). An untabulated Hausman (1978) test indicates that *Δ%EQUITIES* and *ΔERR* are not simultaneously determined. Therefore, as an additional test, I reperform the main tests using OLS instead of 2SLS, since OLS estimates are consistent and efficient in the absence of simultaneity. In the untabulated OLS results, *Δ%EQUITIES_HAT* and *ΔERR_HAT* are not included, since these variables are computed in the first stage of 2SLS. *UNEXLP_ERR* remains significant in the predicted directions. As a caveat, I am not able to rule out the possibility that the Hausman test fails to detect simultaneity even if it exists. Therefore, I use 2SLS in Tables 5 and 6.

[29] As a robustness test, I reperform the tests in Table 5 using the 879 firms in row 15 of Table 3, Panel A and exclude the variables *HORIZON* and *OPER_RISK*. *UNEXPL_ERR* remains statistically significant in the predicted directions.

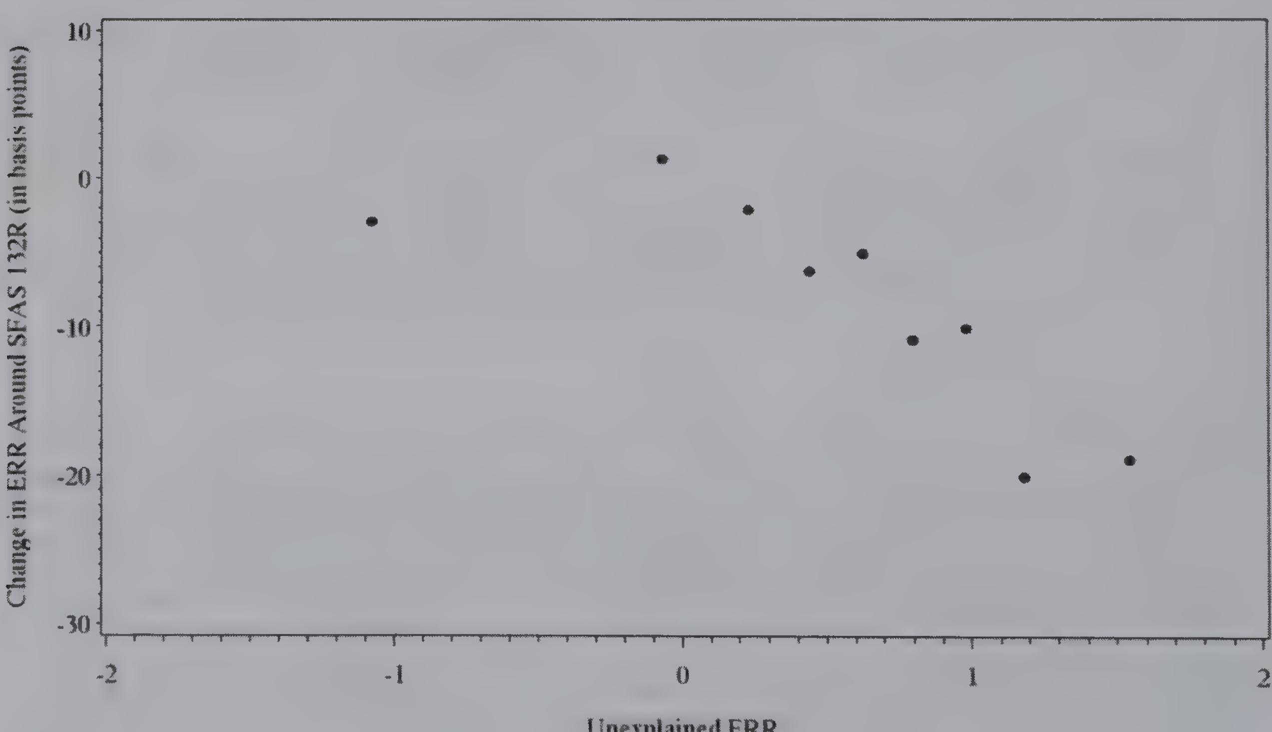
FIGURE 1

Relation between Managerial Responses and Unexplained ERR

Panel A: Change in *%Equities* versus Unexplained ERR (Portfolios of 100 Observations)



Panel B: Change in ERR versus Unexplained ERR (Portfolios of 100 Observations)



The unexplained ERR is the prediction error in the year immediately preceding SFAS 132R computed using parameter estimates from regressing ERRs on asset allocations using the post-SFAS 132R years as an estimation period. Changes in *%EQUITIES* and *ERR* are measured around SFAS 132R. I form portfolios of 100 observations on the unexplained ERR and plot the mean values of the variables for each portfolio. Units on the vertical axis in Panel A are in the change in percent, not the percentage change.

Robustness Test: Potential Confounding Event

As discussed in Section IV, the SEC publicly announced in December 2002 that it would likely challenge firms with ERRs above 9 percent. One alternative explanation is that the firm responses observed in this study are explained not by an accounting rule change but instead by the SEC announcement. To examine this alternative explanation, I develop predictions about the effects of the SEC warning on firm behavior that differ from the effects of SFAS 132R. I expect that firms' adjustments in ERR and pension asset composition in response to SFAS 132R likely apply to all firms whose pre-SFAS 132R ERRs are higher than justified by asset composition. In contrast, I expect that firms' adjustments in ERR and asset composition in response to the SEC's warning in 2002 likely apply only to firms with ERRs greater than the SEC's explicit threshold of 9 percent. Using these differential predictions, I perform two sets of robustness tests.

First, I construct a sample that omits firms with 2002 ERRs greater than the SEC's threshold of 9 percent (comprising 12.6 percent of the sample) to mitigate the possibility that the main results are driven by these firms. Column I in Table 6 reports the results of estimating Equations (2) and (3) after dropping this subsample. The variable of interest for the main tests of H1 and H2 (*UNEXPL_ERR*) remains statistically significant. Stated differently, even when focusing only on firms whose changes in ERRs and asset composition are less likely to be explained by the SEC warning, I still find that firms with pre-SFAS 132R ERRs that are higher than justified by their asset composition respond to SFAS 132R by increasing asset allocation to equities and/or reducing their ERRs.

Second, I isolate the effect of the firm responses for firms whose ERRs exceed the SEC's threshold by creating an indicator variable (*ERR>9%*) set equal to 1 for firms with 2002 ERRs greater than 9 percent, and 0 otherwise. I predict that in response to the SEC's warning, firms with pre-SFAS 132R ERRs greater than the SEC's threshold will reduce ERRs and increase pension allocation to equity securities in 2003 to reduce the likelihood of SEC scrutiny. Column II in Table 6 reports the results of adding this indicator variable to the main regressions. After controlling for the effect of the SEC warning, the variable of interest (*UNEXPL_ERR*) remains statistically significant in the predicted directions.

In summary, the results of the two robustness tests suggest that it is unlikely that changes in firms' ERRs and asset allocations observed in the main tests are explained by the SEC warning in 2002 instead of the accounting rule change under SFAS 132R.

Additional Sensitivity Tests

In this subsection, I discuss the results from a variety of untabulated (independent) sensitivity tests using alternative specifications for the main tests reported in Table 5. First, I add industry indicators to Equations (2) and (3). The R^2 increases to 23.9 percent in Equation (2) and 35.5 percent in Equation (3). The substantial increase in R^2 for ΔERR supports the idea that firms use peer data when choosing ERRs. *UNEXPL_ERR* remains statistically significant. The other control variables remain statistically significant, except that *FIRM_SIZE* and *EQUITY_RET_EXP* lose significance in Equation (3). In the untabulated first-stage regressions, 13 and 30 of the 40 industry indicators are significant for Equations (2) and (3), respectively. Second, I delete all non-December 31 firms.[30] *UNEXPL_ERR* remains statistically significant in the predicted directions. Third, I restrict the tests to a subsample of firms with nonmissing data during all years from 2002 to 2008. The variable of interest *UNEXPL_ERR* remains significantly positive in explaining changes in

---

[30] In tests for the subsample of December 31 firms, I exclude the variables *EQUITY_RET*, *BOND_RET*, *ΔINTEREST_RATE*, and *EQUITY_RET_EXP*, which have the same values for firms with the same year-end date.

TABLE 5

Tests of H1 and H2: Managerial Responses to SFAS 132R
Results of Second-Stage Regressions: Equations (2) and (3)

Panel A: Results of Estimating Equation (2)

$$\Delta\%EQUITIES_i = c0 + c1UNEXPL\_ERR_i + c2\Delta ERR\_HAT_i + \sum Controls_i + u_i$$

Variable	Predicted Sign		Coefficient	t-stat	p-value
Intercept			−0.039	−1.21	0.227
UNEXPL_ERR	+	(H1)	0.030	5.92	<0.001
ΔERR_HAT	+		−0.080	−2.74	0.997
ARR	+		0.246	2.19	0.014
FUNDING	?		0.035	2.27	0.024
HORIZON	+		0.001	0.14	0.446
OPER_RISK	−		−0.003	−0.29	0.384
EQUITY_RET	+		0.218	3.89	0.000
BOND_RET	−		−0.175	−0.36	0.359

n = 768
R^2 = 19.5%
(R^2 from first-stage regression of *Δ%EQUITIES* on instruments: 20.0%)

Panel B: Results of Estimating Equation (3)

$$\Delta ERR_i = d0 + d1UNEXPL\_ERR_i + d2\Delta\%EQUITIES\_HAT_i + \sum Controls_i + u_i$$

Variable	Predicted Sign		Coefficient	t-stat	p-value
Intercept			−0.566	−2.31	0.021
UNEXPL_ERR	−	(H2)	−0.216	−4.06	<0.001
ΔEQUITIES_HAT	+		1.323	1.22	0.112
ARR_AVG3YEAR	+		0.704	1.81	0.036
FIRM_SIZE	−		−0.033	−2.57	0.005
DB_SIZE	−		0.035	2.31	0.990
ΔINTEREST_RATE	+		0.075	1.90	0.029
EQUITY_RET_EXP	+		1.691	1.95	0.026
ERR_LESS_PEER_AVG	−		−0.135	−7.37	<0.001

n = 768
R^2 = 25.1%
(R^2 from first-stage regression of *ΔERR* on instruments: 25.1%)

UNEXPL_ERR is the prediction error in the year immediately preceding SFAS 132R computed using parameter estimates from regressing ERRs on asset allocations and ARR using the post-SFAS 132R years as an estimation period. Please see Table 2, Panel B for the computation of *UNEXPL_ERR*. *FIRM_SIZE* is the log of market capitalization. *DB_SIZE* is the log of the unscaled PBO. See Table 1 for all other variable definitions. *Δ%EQUITIES_HAT* (*ΔERR_HAT*) is the fitted value from a first-stage regression of *Δ%EQUITIES* (*ΔERR*) on a set of instruments. The set of instruments for the first-stage regressions includes: *UNEXPL_ERR*, *ARR*, *ARR_AVG3YEAR*, *FUNDING*, *HORIZON*, *OPER_RISK*, *EQUITY_RET*, *BOND_RET*, *FIRM_SIZE*, *DB_SIZE*, *ΔINTEREST_RATE*, *EQUITY_RET_EXP*, and *ERR_LESS_PEER_AVG*. Reported p-values are one-tailed when there are predicted signs.

TABLE 6

Robustness Tests:
To Rule Out Alternative Explanation that Firm Responses Are Driven by SEC Warning Instead of SFAS 132R
Results of Estimating Modified Versions of Equations (2) and (3)
Column I: Delete Pre-SFAS 132R ERRs Greater than 9 Percent
Column II: Include an Indicator Variable for Pre-SFAS 132R ERRs Greater than 9 Percent

Panel A: Results of Estimating Equation (2)

$$\Delta\%EQUITIES_i = c0 + c1UNEXPL\_ERR_i + c2\Delta ERR\_HAT_i + \sum Controls_i + u_i$$

			Column I: Delete 2002 ERRs > 9%			Column II: Indicator for 2002 ERRs > 9%		
Variable	**Predicted Sign**		**Coefficient**	**t-stat**	**p-value**	**Coefficient**	**t-stat**	**p-value**
Intercept			−0.044	−1.26	0.208	−0.045	−1.37	0.172
UNEXPL_ERR	+	(H1)	0.027	4.90	<0.001	0.030	5.69	<0.001
ΔERR_HAT	+		−0.073	−2.49	0.993	−0.082	−2.84	0.998
ARR	+		0.253	2.00	0.023	0.266	2.31	0.011
FUNDING	?		0.057	2.75	0.006	0.035	2.20	0.028
HORIZON	+		0.005	0.83	0.204	0.002	0.31	0.380
OPER_RISK	−		−0.002	−0.21	0.416	−0.003	−0.32	0.376
EQUITY_RET	+		0.207	3.42	0.000	0.225	3.95	<0.001
BOND_RET	−		−0.240	−0.47	0.319	−0.151	−0.31	0.379
ERR>9%	+					−0.010	−1.01	0.843
n			671			768		
R^2			18.9%			19.6%		

Panel B: Results of Estimating Equation (3)

$$\Delta ERR_i = d0 + d1UNEXPL\_ERR_i + d2\Delta\%EQUITIES\_HAT_i + \sum Controls_i + u_i$$

			Column I: Delete 2002 ERRs > 9%			Column III: Indicator for 2002 ERRs > 9%		
Variable	**Predicted Sign**		**Coefficient**	**t-stat**	**p-value**	**Coefficient**	**t-stat**	**p-value**
Intercept			−0.562	−2.44	0.015	−0.537	−2.26	0.025
UNEXPL_ERR	−	(H2)	−0.230	−4.86	<0.001	−0.213	−4.07	<0.001
ΔEQUITIES_HAT	+		1.199	1.23	0.110	1.028	0.97	0.165
ARR_AVG3YEAR	+		0.968	2.27	0.012	0.963	2.50	0.006
FIRM_SIZE	−		−0.030	−2.21	0.014	−0.027	−2.16	0.015
DB_SIZE	−		0.035	2.14	0.984	0.031	2.07	0.981
ΔINTEREST_RATE	+		0.099	2.47	0.007	0.085	2.17	0.015
EQUITY_RET_EXP	+		1.573	1.90	0.029	1.520	1.80	0.036
ERR_LESS_PEER_AVG	−		−0.144	−7.73	<0.001	−0.135	−7.55	<0.001
ERR>9%	−					0.004	0.12	0.547
n			671			768		
R^2			28.3%			27.0%		

(continued on next page)

TABLE 6 (continued)

The purpose of this table is to attempt to isolate the effect of the SEC warning that it would challenge ERRs above 9 percent. In Column I, I delete all observations with pre-SFAS 132R ERRs greater than 9 percent. In Column II, I include the variable *ERR>9%*, which is an indicator variable set equal to 1 if the pre-SFAS 132R ERR is greater than 9 percent, and 0 otherwise. All other variables are the same as in Table 5. Reported p-values are one-tailed when there are predicted signs.

equity allocation after SFAS 132R, but is marginally negative for changes in ERR. Fourth, I delete firms with flat pension schemes where the absolute difference between PBO and ABO is less than 5 percent of PBO (where both PBO and ABO are nonmissing). *UNEXPL_ERR* remains statistically significant in the predicted directions. Fifth, I restrict the tests to a subsample of firms with overfunded pensions in the first year of SFAS 132R. *UNEXPL_ERR* remains significantly positive in explaining changes in ERR but is not significant at explaining changes in equity allocation, possibly due to the small sample size of 177 firms with overfunded pensions. Sixth, I restrict the tests to a subsample of firms with underfunded pensions in the first year of SFAS 132R. *UNEXPL_ERR* remains statistically significant in the predicted directions. Seventh, I estimate Equations (2) and (3) in the years 2004 to 2008, instead of 2003. In years other than the first year of SFAS 132R, the coefficient for *UNEXPL_ERR* is not significant. The finding that the effects predicted by H1 and H2 do not arise in different time periods is consistent with the interpretation that SFAS 132R is responsible for the managerial behavior documented in the first year of SFAS 132R.

Finally, using three alternative specifications, I address the possibility that the error term in Equation (3) is not normally distributed.[31] In the first of three specifications, I define a variable that equals 1 if $\Delta ERR < 0$, 2 if $\Delta ERR = 0$, and 3 if $\Delta ERR > 0$. I use this discrete ordered ΔERR variable in the second-stage, where the hypothesis testing occurs. I use OLS to estimate both the first and second stages. In the second specification, I use the same discrete ordered ΔERR variable but estimate the second-stage regression using an ordered logistic regression for ΔERR. In the third specification, I transform both dependent variables. I define an indicator variable that equals 1 if $\Delta ERR < 0$, and 0 otherwise. I define another indicator variable that equals 1 if $\Delta\%EQUITIES > 0$, and 0 otherwise. Results of this specification can be interpreted as a choice study.[32] For these three specifications, the R^2 generally decreases compared to Table 5, and many of the control variables lose significance, but *UNEXPL_ERR* remains significant in the predicted directions. Overall, the results in this subsection support the idea that the main results are robust to various specifications.

Supplemental Analysis Using IRS Form 5500

As discussed in Section III, my maintained assumption is that firms react to SFAS 132R despite the fact that, even before SFAS 132R, the IRS required firms with pensions plans covering

[31] Many firms do not alter their ERRs from year to year, such that many firms have values of zero for the continuous ΔERR variable. OLS assumes the error term is normally distributed. A distribution of the dependent variable with a concentration at the same value (such as that observed for ΔERR with a concentration at zero) could potentially result in non-normal errors. If the errors are not normally distributed, then the parameter estimates are consistent but not efficient (Kennedy 1992), thereby affecting significance tests in hypothesis testing. A common technique to address non-normal error terms is to transform the dependent variable.

[32] For ease of interpretation as a choice study, I use these indicator variables and logistic regressions in both the first and second stages.

more than 100 participants to report asset composition on Form 5500. In this section, I investigate whether the SFAS 132R disclosures provide incremental information beyond Form 5500.[33] Detailed results are available upon request.

I perform several analyses using Form 5500 data, which offer three major takeaways. First, based on untabulated descriptive statistics, most firms do not report data for the individual categories on Form 5500 (i.e., most of the more detailed categories have means and medians of zero). Firms typically report assets in broader categories, such as trusts. Second, based on untabulated regressions of the ERR on Form 5500 asset categories, it is unlikely that financial statement users would be able to judge the reasonableness of the ERR based on the asset allocations reported on Form 5500. Third, I reperform the main tests in Table 5, except that I compute *Δ%EQUITIES* and *UNEXPL_ERR* using asset allocations from Form 5500. *UNEXPL_ERR* is not significant in the predicted directions. In summary, while Form 5500 provides useful information about pension asset allocation, SFAS 132R disclosures appear to provide asset composition information that is more accessible and visible.

VI. CONCLUSION

This study examines whether firms alter their behavior in response to changes in accounting standards that mandate new financial statement disclosures. The primary incremental contribution of this study is that firms alter their behavior in response to changes in standards relating merely to *disclosure*—as opposed to recognition or measurement. Disclosure requirements arguably create less powerful incentives to alter firm decisions than recognition requirements. Finding that new disclosure requirements are enough to induce meaningful changes in firm behavior is compelling evidence in support of the idea that firms do alter their behavior in response to changes in accounting standards.

I examine my research question in the context of defined benefit pensions, which offers an interesting setting with large and economically meaningful effects to study the consequences of FASB accounting standards. In particular, I examine whether firms take actions to more closely align the expected rate of return (ERR) with asset allocation following the mandated asset composition disclosure. Findings suggest that, while it appears that SFAS 132R achieved one of its stated objectives, there were also economic consequences following the issuance of this accounting standard. I find that firms that used upward-biased ERRs prior to SFAS 132R responded to the new standard by increasing asset allocation to equities and/or by reducing ERRs.

Increased allocation to equities in defined benefit pension plans is a particularly interesting economic consequence for two reasons. First, at the firm level, a shift in pension assets toward equities can contribute to a risk mismatch between pension assets and pension liabilities, because pension assets invested in equities have durations and risk characteristics different from those of pension liabilities (Merton 2006). If there are large stock market losses, such as those in 2008, then pension plans can become severely underfunded. A Mercer Consulting (2009) study estimates that stock market losses left the pension plans of the S&P 1500 firms underfunded by $409 billion as of

---

[33] I obtain machine-readable Form 5500 data from the Center for Retirement Research at Boston College. The IRS collects the Form 5500 data at the plan level, not the firm level. Following prior studies, I aggregate the data to the firm level using the employer identification number (EIN), which is the firm-level equivalent of a social security number. I supplement this process with a manual matching based on firm name and state of incorporation. I then merge the Form 5500 data to Compustat using EINs. As discussed in prior studies, there are difficulties merging Form 5500 with Compustat (Gron and Madrian 2004). The entity that sponsors the pension could be a controlled subsidiary choosing to file taxes separately from the parent and thus has a separate EIN from the parent, while still remaining consolidated with the parent for financial reporting purposes. As a result, there is a lack of a unique common identifier between Form 5500 and Compustat. I follow the suggestions provided in Gron and Madrian (2004).

December 2008. Thus, firms that increase equity allocation in order to report a lower pension expense can be exposing their shareholders to large risks.

Second, my findings suggest that the economic consequences of FASB standards can contribute to increased correlation of stock price performance among firms within an economy and aggregate pension underfunding (i.e., when pension obligations exceed pension assets). Shareholders of a pension-sponsoring firm are the residual claimants on pension-funding deficits or surpluses. If a negative price shock reduces the value of the pension plan assets, then the value of the pension-sponsoring firm is reduced because shareholders bear the costs of pension deficits through increased future contributions to the pension. To the extent that firms cross-hold each other's equities in their defined benefit pensions, pension-funding levels and firm values become more highly correlated and interdependent across firms. In summary, firms' responses to FASB standards can result in actions that contribute to increased risk exposure at both the firm level and the macroeconomic level.

REFERENCES

Accounting Standards Board (ASB). 2000. *Retirement Benefits*. Financial Reporting Standards No. 17. London, U.K.: ASB.

Amir, E., and A. Ziv. 1997. Recognition, disclosure, or delay: Timing the adoption of SFAS No. 106. *Journal of Accounting Research* 35 (1): 61–81.

Amir, E., and S. Benartzi. 1998. The expected rate of return on pension fund and asset allocation as predictors of portfolio performance. *The Accounting Review* 73 (3): 335–352.

Amir, E., Y. Guan, and D. Oswald. 2010. The effect of pension accounting on corporate pension asset allocation. *Review of Accounting Studies* 15 (2): 345–366.

Barth, M. 1991. Relative measurement errors among alternative pension asset and liability measures. *The Accounting Review* 66 (3): 433–463.

Beatty, A. 1995. Fair value accounting: How fair is it? *St. Louis Federal Review*: 25–39.

Bens, D., and S. Monahan. 2008. Altering investment decisions to manage financial reporting outcomes: Asset-backed commercial paper conduits and FIN 46. *Journal of Accounting Research* 46 (5): 1017–1055.

Bergstresser, D., M. Desai, and J. Ruah. 2006. Earnings manipulations, pension assumptions, and managerial investment decisions. *The Quarterly Journal of Economics* 121 (1): 157–195.

Biddle, G., and G. Seow. 1991. The estimation and determinants of association between returns and earnings: Evidence from cross-industry comparisons. *Journal of Accounting, Auditing, and Finance* 6 (2): 183–232.

Black, F. 1980. The tax consequences of long-run pension policy. *Financial Analysts Journal* 36 (4): 21–28.

Bloomberg Businessweek Magazine. 2004. Pumped-up pension plays. (October 24).

Bodie, Z., J. Light, R. Morck, and R. Taggart. 1987. Funding and asset allocation in corporate pension plans: An empirical investigation. In *Issues in Pension Economics*, edited by Z. Bodie, J. B. Shoven, and D. A. Wise, 15–48. Cambridge, MA: University of Chicago Press.

Bruno, M. 2005. SEC probe another headache for beleaguered GM. *Pensions and Investments* (October 31).

Choudhary, P., S. Rajgopal, and M. Venkatachalam. 2008. Accelerated vesting of stock options in anticipation of SFAS 123R. *Journal of Accounting Research* 47 (1): 105–146.

Consolidated Edison. 2002. *2002 Annual Report*. Available at: http://media.corporate-ir.net/media_files/nys/ed/reports/Con_Edison_2002_Annual_Report.pdf

Financial Accounting Standards Board (FASB). 1985. *Employers' Accounting for Pensions*. Statement of Financial Accounting Standards 87. Norwalk, CT: FASB.

Financial Accounting Standards Board (FASB). 2003. *Employers' Disclosures About Pensions and Other Postretirement Benefits*. Statement of Financial Accounting Standards 132R. Norwalk, CT: FASB.

Franzoni, F., and J. Marin. 2006. Pension plan funding and stock market efficiency. *Journal of Finance* 61 (2): 921–956.

Friedman, D. 1983. Pension funding, pension asset allocation, and corporate finance: Evidence from individual company data. In *Financial Aspects of the United States Pension System*, edited by Z. Bodie, and J. B. Shoven, 107–152. Chicago, IL: University of Chicago Press.

Graham, J. 1996. Proxies for the corporate marginal tax rate. *Journal of Financial Economics* 42 (2): 187–221.

Graham, J., C. Harvey, and S. Rajgopal. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40 (1-3): 3–73.

Graham, J., and L. Mills. 2008. Using tax data to simulate corporate marginal tax rates. *Journal of Accounting and Economics* 46 (2-3): 366–388.

Gron, A., and B. Madrian. 2004. *Matching IRS Form 5500 Filings with Compustat and CRSP*. Working paper, Harvard University.

Hann, R., F. Heflin, and K. Subramanyam. 2007. Fair-value pension accounting. *Journal of Accounting and Economics* 44 (3): 328–358.

Harrison, J., and W. Sharpe. 1983. Optimal funding and asset allocation rules for defined benefit pension plans. In *Financial Aspects of the United States Pension System*, edited by Z. Bodie, and J. Shoven, 91–106. Chicago, IL: University of Chicago Press.

Hausman, J. 1978. Specification tests in econometrics. *Econometrica* 46 (6): 1251–1271.

Horwitz, B., and R. Kolodny. 1980. The economic effects of involuntary uniformity in the financial reporting of research and development expenditures. *Journal of Accounting Research* 18: 38–74.

Imhoff, E., and J. Thomas. 1988. Economic consequences of accounting standards: The lease disclosure rule change. *Journal of Accounting and Economics* 10 (4): 277–310.

Jin, L., R. Merton, and Z. Bodie. 2006. Do a firm's equity returns reflect the risk of its pension plan? *Journal of Financial Economics* 81 (1): 1–26.

Kennedy, P. 1992. *A Guide to Econometrics*. Cambridge, MA: MIT Press.

Libby, R., M. Nelson, and J. Hunton. 2006. Recognition v. disclosure, auditor tolerance for misstatement, and the reliability of stock-based compensation and lease information. *Journal of Accounting Research* 44 (3): 533–560.

Mercer Consulting. 2009. Quarterly pensions update. January 2009. Available at: http://www.mercer.com

Merton, R. 2006. Allocating shareholder capital to pension plans. *Journal of Applied Corporate Finance* 18 (1): 15–24.

Mittelstaedt, H., W. Nichols, and P. Regier. 1995. SFAS No. 106 and benefit reduction decisions in employer-sponsored retiree health care plans. *The Accounting Review* 70 (4): 535–556.

Picconi, M. 2006. The perils of pensions: Does pension accounting lead investors and analysts astray? *The Accounting Review* 81 (4): 925–955.

Revell, J. 2002. Beware the pension monster. *Fortune* (December 3).

Stuart, A. 2005. Death to smoothing. *CFO Magazine* (February 22).

Tepper, I. 1981. Taxation and corporate pension policy. *Journal of Finance* 36 (1): 1–13.

Zhang, H. 2009. Effect of derivative accounting rules on corporate risk-management behavior. *Journal of Accounting and Economics* 47 (3): 244–264.

Zion, D., and B. Carache. 2002. *The Magic of Pension Accounting, Part I*. Boston, MA: Credit Suisse First Boston.

Zion, D., and B. Carache. 2005. *The Magic of Pension Accounting, Part III*. Boston, MA: Credit Suisse First Boston.

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 429–462

American Accounting Association
DOI: 10.2308/accr-50317

How Much Does IFRS Cost? IFRS Adoption and Audit Fees

Emmanuel T. De George
University of Michigan

Colin B. Ferguson
Nasser A. Spear
The University of Melbourne

ABSTRACT: This study provides evidence of a directly observable and significant cost of International Financial Reporting Standards (IFRS) adoption, by examining the fees incurred by firms for the statutory audit of their financial statements at the time of transition. Using a comprehensive dataset of all publicly traded Australian companies, we quantify an economy-wide increase in the mean level of audit costs of 23 percent in the year of IFRS transition. We estimate an abnormal IFRS-related increase in audit costs in excess of 8 percent, beyond the normal yearly fee increases in the pre-IFRS period. Further analysis provides evidence that small firms incur disproportionately higher IFRS-related audit fees. We then survey auditors to construct a firm-specific measure of IFRS audit complexity. Empirical findings suggest that firms with greater exposure to audit complexity exhibit greater increases in compliance costs for the transition to IFRS. Given the renewed debate about whether the Securities and Exchange Commission (SEC) should mandate IFRS for U.S. firms, our results are of timely importance.

Keywords: *IFRS; audit fees; international harmonization; IFRS costs.*

Data Availability: *Data are publicly available from the sources identified in the paper. Survey response data are available from the authors upon request.*

I. INTRODUCTION

Regulators and standard setters claim that International Financial Reporting Standards (IFRS) enhance the comparability and quality of financial reporting. However, the true returns to IFRS adoption should be evaluated by trading off the costs of transition and any

We thank Paul Coram, Larelle Chapple, John Harry Evans III (senior editor), Gerry Gallery, Clive Lennox, Russell Lundholm, Greg Miller, Michael Minnis, Venky Nagar, Terry O'Keefe, Mo Kahn, Denis Oswald, Nemit Shroff, Cathy Shakespeare, Wayne Thomas (editor), Hal White, Gwen Yu, Ian Zimmer, and two anonymous referees, along with seminar participants at the AFAANZ Conference, Deakin University, Queensland University of Technology, University of Melbourne, University of Michigan, and University of Adelaide for their useful comments. All errors and omissions are our own.

Editor's note: Accepted by Wayne Thomas.

Submitted: September 2009
Accepted: September 2012
Published Online: October 2012

recurring costs of reporting against the recurring benefits of comparability and increased reporting quality (Hail et al. 2010). To date, most empirical research into IFRS has not explicitly separated the gross costs and benefits of adoption, but rather, has focused on the net impact to capital markets and financial reporting outcomes pre- and post-transition in assessing its merits (e.g., Goodwin and Ahmed 2006; Armstrong et al. 2010; Daske et al. 2008). This empirical approach has provided valuable, yet mixed, conclusions: the *net* benefits of IFRS adoption vary significantly across firms, industries, and countries (Daske et al. 2008).[1] Such variation calls for a better understanding of the specific nature of the costs and benefits to determine the economic trade-off (see Hail et al. 2010). Given the impending decision and renewed debate surrounding whether the Securities and Exchange Commission (SEC) should mandate the use of IFRS in the U.S. market, understanding the costs of IFRS implementation is of timely importance. This study quantifies the directly observable and significant cost of IFRS compliance by examining the fees incurred by firms for the statutory audit of their financial statements.[2]

Our focus on audit costs is motivated by the fact that the pervasive nature of IFRS adoption is likely to have a profound impact on a firm's financial reporting costs (Hail et al. 2010). Firms currently subject to IFRS have already raised concerns over increasing preparation and certification costs (Jermakowicz and Gornik-Tomaszewski 2006). Survey evidence following the mandatory adoption of IFRS within the European Union (EU) suggests per-firm estimates of transition costs in the order of 0.31 percent of total sales for small firms; i.e., less than $700 million in sales, and up to 0.05 percent for large firms (Institute of Chartered Accountants in England and Wales [ICAEW] 2007). Extrapolating this to a U.S. adoption, Hail et al. (2010) place the expected aggregate IFRS transition cost at approximately eight billion dollars. However, a recent survey carried out by Accenture found that U.S. executives actually expect to pay *more* than their EU counterparts, estimating between 0.1–0.7 percent of annual revenues, depending on firm size (Johnson 2009). Large-scale empirical examination of compliance costs allows us to assess the validity of these conjectures, and more accurately predict the costs that may be incurred for a U.S. adoption of IFRS.

From an empirical standpoint, audit fees represent a direct, observable, and measurable cash outflow that incorporates significant changes in accounting regulations. Examination of audit fees allows us to draw insights into the effects of IFRS on the audit function. Given the significance of the underlying institutional elements in place in IFRS-adopting jurisdictions, auditors are likely to play a key role in shaping future reporting practices related to IFRS (Ball 2006).

The Australian setting offers four advantages in addressing our research question: (1) Australia, a member of the "G4+1" (United States, United Kingdom, Australia, Canada, and New Zealand) common law countries, has comparatively deep markets, developed shareholders' rights, auditing services, and other monitoring systems similar to those in the U.S. (Ball 2006), which allows some generalization of the costs expected with a U.S. adoption; (2) the adoption of IFRS has a clear starting point (financial periods beginning on or after January 1, 2005) without a staggered adoption process, and virtually no cases of voluntary early adoption (less than 1 percent); (3) availability of a history of high-quality and comparable audit disclosures and fee data; and, perhaps most importantly, (4) the blanket Australian adoption of both public and private companies prevented firms from "opting-out" and, thus, mitigates potential self-selection bias from voluntary delistings.

---

[1] For instance, value relevance of IFRS-compliant earnings is found to increase in Germany, relative to local GAAP (Jermakowicz et al. 2007), vary significantly in the U.K. (Horton and Serafeim 2010), and show no improvement in Spain (Callao et al. 2007). The observed capital market net benefits, such as cost of capital and liquidity effects, exhibit significant variation across countries, industries, and firms (Daske et al. 2008).

[2] While we motivate the current study with references to the costs and benefits underpinning the IFRS transition effects on firms, we are silent on any associated benefits of adoption. We view the primary contribution of this study as the understanding of the cost associated with IFRS adoption in order to shed light on the existing variation documented in "net benefit" studies of IFRS.

Employing a comprehensive sample of all publicly traded companies on the Australian Stock Exchange for the period 2002–2006, we estimate a 23 percent increase in the average level of audit fees in the year of IFRS adoption. Further examination of the changes in audit fees finds an *abnormal* increase of 8 percent in the year of IFRS adoption, beyond normal yearly increases in fees.[3] We then incorporate predicted cross-sectional variation in the observed audit fee increases based on *IFRS-exposure*, defined as the magnitude of net IFRS adjustments made to total equity. We find a positive and significant relation between fee increases and *IFRS-exposure*. Results suggest the observed increase in audit fees includes a fixed component of approximately 10 percent exhibited by all firms, and an incremental *complexity* component of 16 percent (17 percent) by firms with positive (negative) IFRS adjustments. Firms that reported the largest IFRS transitional adjustments, i.e., greater than 6 percent of total assets, incurred the greatest increase of over 30 percent in audit costs. These fees were almost 50 percent larger than fees for firms that reported little or no IFRS adjustments. These results remain robust to alternative estimation methodologies and model specifications. We also consider the effects of IFRS adoption on the underlying measurement of our control variables, to ensure our results are not a spurious artifact of pooling pre- and post-IFRS observations. Our results are consistent with IFRS transition costs being a function of firm-specific exposure to IFRS and a fixed economy-wide switching cost.

We also provide evidence that IFRS are onerous for small firms. It has been argued that IFRS reporting costs are "likely to have a fixed component, making certain reports or disclosures particularly burdensome for smaller firms" (Hail et al. 2010, 13). A recent SEC roundtable of CEOs from small public companies commented that a U.S. adoption of IFRS would be "painful and costly," with little or no return (Segarra 2011). Executives at small U.S. firms expect a disproportionately higher cost of IFRS transition relative to large firms, predicting transitional costs in the order of 0.731 percent of annual revenues (Johnson 2009). While large firms tend to have adequate technical knowledge and internal resources to cope with a reporting regime change, small firms are less likely to have this depth of technical knowledge, and find such transitions more challenging (Kittney and Buffini 2006). Our results find the estimated percentage increase in audit fees associated with IFRS adoption is approximately 1.6 times greater for small firms. Moreover, we find that small firms exhibit a significant fixed-cost component to their fee increase in the year of IFRS adoption. This increase is in addition to their incremental fee increase explained by their IFRS adjustments. In contrast, the audit fee increases incurred by large firms are primarily explained by the magnitude of their IFRS adjustments, with no discernible fixed-cost component.

We then provide insight concerning which aspects of IFRS are most costly. To do this, we survey Big 4 auditors on specific IFRS requirements, in terms of incremental audit effort, complexity, and risks relative to previous Generally Accepted Accounting Principles (GAAP). The auditors' responses indicate that IFRS requirements for financial instruments, share-based payments, intangible assets, and income taxes warrant considerable effort and entail the most audit complexity. Incorporating a novel variable that captures a firm's aggregate exposure to these specific IFRS requirements, *IFRS_score*, we find firms' fee increases are positively related to their exposure to these key IFRS requirements.

Our findings contribute to the literature in several ways. We document and quantify audit fees as a significant cost associated with mandatory adoption of IFRS. Our empirical results concerning small firms also inform discussions on the appropriateness of mandated IFRS regulations for small to mid-sized entities. In addition, our study is novel in that we obtain expert opinions from professional auditors of the impact of IFRS on the audit function. Specifically, we isolate the

[3] We use the term "*abnormal*" throughout the paper to refer to the estimated increases in audit fees beyond the natural trend in yearly fee increases occurring in the pre-IFRS period.

specific financial statement items and requirements that require the greatest audit effort during IFRS adoption. Given the emphasis of auditing in the proposed roadmap for U.S. convergence to IFRS (Jamal et al. 2009; SEC 2008), our results will be of particular importance to U.S. firms, auditors, and regulators. On a broader level, we document variation in the costs of IFRS adoption that provides insight into the significant variation in the net benefit observed in previous studies. For instance, we find evidence that certain IFRS requirements have higher compliance costs than others, which explains the significant heterogeneity in compliance costs during the adoption of IFRS.

Section II discusses the nature of IFRS adoption costs and the impact on the audit function. Section III describes our data. Section IV presents and discusses our analysis of the impact of mandatory IFRS adoption on the audit fees incurred by all firms in the Australian economy. Section V presents our results regarding the differential costs of IFRS adoption on small and large firms. Section VI presents survey and empirical results from professional auditors regarding the specific IFRS regulations that generate the most audit costs. Section VII concludes the paper.

II. IFRS ADOPTION COSTS AND THE AUDIT FUNCTION

We view IFRS adoption as costly to firms because of the greater effort, knowledge, and information systems needed to implement the new standards, and the additional effort needed to manage the risk of material misstatements appearing in IFRS-compliant financial statements. Firms currently subject to IFRS have raised concerns over the significant preparation and certification costs required by this change in reporting regime (see Jermakowicz and Gornik-Tomaszewski 2006). U.S. executives estimate significant firm-specific IFRS transitional costs between 0.1 to 0.7 percent of annual revenues (Johnson 2009). We examine the issue of IFRS compliance costs by focusing on audit fees at the time of adoption.

Prior studies that have examined the impact of mandated shifts in accounting and corporate governance regulation on the audit function suggest these costs are significant.[4] They attribute these costs primarily to increased audit effort and the associated increased audit risk around the time of mandated shifts in regulation (Ghosh and Pawlewicz 2009; Charles et al. 2010; Pricewaterhouse-Coopers 2008).

In the case of IFRS adoption, we expect the increased audit effort to be driven by two main factors. First, auditors will exert additional effort to become knowledgeable about the new standards so they can assess whether these standards have been appropriately implemented. Auditors are likely to seek to recover the cost of this increased effort through increases in audit fees in the year of IFRS adoption. Recurring costs will increase if auditors assess that a continued higher effort is required by the IFRS reporting regime relative to existing local GAAP. In many cases, IFRS are more detailed and involve increased footnote disclosure, requiring the auditor to certify financial information of a differing nature than that previously reported (Webb 2006; Ernst & Young 2005). An example of internal transition/learning costs expended by auditors is the development of IFRS accreditation programs for audit staff, as articulated by KPMG Australia national audit partner John

[4] We are aware of three studies that document country-specific empirical relations between IFRS adoption and audit fees; however, results are mixed. Griffin et al. (2009) document an increase in the mean level of audit fees and residual audit fees in New Zealand, while Schadewitz and Vieru (2008) find only a weak relation between audit fees and IFRS adoption using Finnish data. On a broader level, a recent working paper proposes that increases in audit costs post-IFRS are attributable to increases in audit task complexity, and varied with the strength of a country's legal regime (Kim et al. 2010). Using a comprehensive sample of a developed economy, employing a differing research design, and obtaining survey opinions from auditors, we are able to more fully examine the effects of IFRS adoption on audit fees. We believe the benefits of this rich setting allow more complete and generalizable results for future adopting countries.

Teer, who spoke of "the need to manage additional complexity in the audit" as a result of the global standards (Lynch and Fabro 2005).

Second, auditors are likely to increase audit effort to manage the risk of IFRS adoption in terms of (1) the chances of the financial statements being materially misstated, and (2) the litigation risk (i.e., the probability of incurring liability payments or loss of reputation capital) associated with the consequences of materially misstated financial statements. A primary source of this risk comes from IFRS' emphasis on fair value. The increased reliance on fair value measurement on a number of key account balances will result in more professional judgment, discretion, and subjectivity in the financial reporting process. This increases the potential occurrence of a reporting error and, ultimately, audit failure (Love and Eickemeyer 2009; Lin and Yen 2009). This reduced emphasis on verifiability as a key concept in guiding the development of new accounting standards has led to less prescriptive guidance and increased subjectivity in accounting measurement (Jamal et al. 2009). Overall, there is a view that the principles-based IFRS standards will exacerbate litigation costs, while depriving auditors of their ability to demonstrate compliance with specific guidelines and established rules in defense of an audit failure (Diehl 2010).

Because IFRS require more detailed disclosure than prior GAAP, auditors are now certifying more financial information that includes management's subjective forecasts and assessments of assets and liabilities. For example, the reporting requirements for accounting hedges call for companies to undertake and document detailed tests of hedge effectiveness, and provide significantly more disclosure on assumptions underlying this analysis. Moreover, the IFRS provisions relating to share-based payments require substantial disclosure as to the nature and method of executive compensation plans, along with detailed information on inputs of fair value calculations. It has been reported that first-time IFRS-compliant annual reports are up to 60 percent longer than previous annual reports (Webb 2006; Ernst & Young 2005).

The general uncertainty surrounding IFRS adoption also contributes to the increased compliance costs faced by firms. Uncertainty in the financial reporting environment increases *ex post* investor scrutiny over the relatively new IFRS financial statements, increasing the likelihood of costly shareholder litigation. To protect their reputation capital, auditors increase audit effort and/or client risk assessments (Clarkson et al. 2003), which are likely to result in increases in audit fees. Overall, we expect to observe increased audit fees associated with the adoption of IFRS attributable to increased audit effort, increased investment in audit resources, and an increased audit risk premium.

III. DATA SOURCES

For financial periods beginning on or after January 1, 2005, all Australian public and private firms were required to prepare IFRS-compliant financial statements. Our sample consists of all companies publicly traded on the Australian Stock Exchange (ASX) that adopted IFRS from January 1, 2005, and have sufficient available data for our specifications. We require four years of preceding audit fee data in our *pre*-IFRS period, auditor information, and IFRS transition information, which we hand-collect from publicly available annual reports. In addition to audit fee data and IFRS transition information, we collect data on the number of local and foreign subsidiaries directly from publicly available reports. We collect other financial information relating to the control variables employed in the audit fee regression model from Aspect Huntley's Fin Analysis database for the relevant years.[5]

[5] Fin Analysis is a web-based database that holds PDF versions of all published financial statements filed with the ASX, along with myriad other financial and operating information on all ASX-listed firms.

There are 1,724 listed companies on the ASX that filed first-time IFRS-compliant statements beginning on December 31, 2005. We identify 1,225 firms listed for the full five years of our sample period, and 1,071 of these that issued annual financial statements with valid audit fee disclosures each year. We then remove 101 firms with missing financial and operating data needed for computation of control variables, eight that were preparing IFRS-compliant financial statements prior to the mandated adoption date, 41 firms where we could not identify consistent IFRS adjustments from the note disclosures, and 14 firms engaged in commercial banking activities. This leaves a sample of 907 firms (4,535 firm-years) from 2002 through 2006 for our primary analysis, accounting for approximately 79.6 percent of the total ASX market capitalization as of June 30, 2006.

Table 1 reports the distribution of GICS Industry codes, both two-digit macro-level industry sectors and four-digit industry groups. Consistent with the make-up of the Australian stock market, 253 sample firms come from materials industries, such as construction materials, metals firms, and mineral and mining companies, and 93 are diversified financial firms. To ensure our results are not driven by individual industries, we follow the approach of prior studies (e.g., Ferguson et al. 2003) by dropping all firms within each two-digit macro-level GICS sector and rerunning our main results, as described in later sections.

TABLE 1

Distribution of Firms across Industries

GICS Sector (Two-Digit)	Number of Firms in Sector	GICS Industry Groups (Four-Digit)	Number of Firms in Group
Energy	78	Energy	78
Materials	253	Materials	253
		Capital Goods	52
		Commercial and Professional Services	32
Industrials	98	Transportation	14
		Automobiles and Components	8
		Consumer Durables and Apparel	18
		Consumer Services	24
		Media	29
Consumer Discretionary	98	Retailing	19
		Food and Staples Retailing	6
Consumer Supplies	32	Food, Beverage and Tobacco	26
		Health Care Equipment and Services	45
Health Care	87	Pharmaceuticals and Biotechnology	42
		Diversified Financials	93
		Insurance	6
Financials	157	Real Estate	58
		Software and Services	58
		Technology Hardware	19
Information Technology	78	Semiconductors	1
Telecommunications	14	Telecommunications	14
Utilities	12	Utilities	12
Total	907		907

This table shows the distribution of firms within each two-digit macro-level GICS sector and four-digit Industry Group.

IV. IFRS ADOPTION AND ECONOMY-WIDE TRANSITION COSTS

Levels Analysis

We begin with levels analysis, estimating a variant of the traditional audit fee regression model (Simunic 1980; Ferguson et al. 2003; Craswell et al. 1995), controlling for systematic differences in client size, audit complexity, auditor-client risk sharing, and other changes in the financial reporting environment. We then include a number of experimental variables, including adoption dummies and IFRS exposure measures, to capture any increase in audit fees in the year of adoption, relative to the pre-IFRS period, as outlined below:

$$\begin{aligned} LogAF_{it} = {} & \beta_0 + \beta_1 LogAsset_{it} + \beta_2 LogNAS_{it} + \beta_3 Rec_{it} + \beta_4 Inv_{it} + \beta_5 Accr_{it} + \beta_6 Quick_{it} \\ & + \beta_7 Debt_{it} + \beta_8 ROA_{it} + \beta_9 Forsubs_{it} + \beta_{10} Totsubs_{it} + \beta_{11} loss_{it} + \beta_{12} US\_List_{it} \\ & + \beta_{13} Opinion_{it} + \beta_{14} Big4_{it} + \beta_{15} YE_{it} + \gamma_j IFRS\_Var_{it} + \varepsilon_{it}, \end{aligned} \tag{1}$$

where subscripts refer to firm i in year t, for variables defined as follows:

LogAF = natural log of total audit fees paid to external auditors;
LogAsset = natural log of total assets;
LogNAS = natural log of total nonaudit service fees paid to external auditors;
Rec = ratio of total receivables to ending total assets;
Inv = ratio of total inventory to ending total assets;
Accr = absolute value of accruals (computed as difference between net income and cash flow from operations) scaled by ending total assets;
Quick = ratio of current assets to current liabilities;
Debt = ratio of long-term debt to ending total assets;
ROA = ratio of net profit after tax to ending total assets;
Forsubs = natural log of 1 plus the number of foreign subsidiaries;
Totsubs = natural log of 1 plus the number of total subsidiaries;
US_List = indicator variable equal to 1 if the firm is dual-listed on a U.S. exchange and, hence, subject to SEC enforcement, otherwise equal to 0;
Loss = indicator variable equal to 1 if the firm reported a loss in the current year, and a profit in the previous year, otherwise equal to 0;
Opinion = indicator variable equal to 1 if the firm was issued with a modified opinion in the current year, otherwise equal to 0;
Big4 = indicator variable equal to 1 if the firm is audited by a Big 4 auditor in the current year, otherwise equal to 0;
YE = indicator variable equal to 1 for firms with a June 30 year-end, otherwise equal to 0; and
IFRS_Var = experimental variables capturing IFRS effect on mean audit fees in the period of IFRS adoption, relative to pre-IFRS period. Defined in detail below.

We adjust all dollar-denominated variables for inflation using 2002 as a base year. We control for auditee size (*LogAsset*), audit complexity (*Accr*, *Rec*, *Inv*, *Forsubs*, *Totsubs*), loss exposure (*Quick*, *Debt*), peak-season audits (*YE*), and auditor-client risk sharing (i.e., the likelihood that litigation losses are borne by the auditor) (*Opinion*, *ROA*, *Loss*). In addition, we control for Big 4 versus non-Big 4 auditors (*Big4*); aggregate nonaudit service fees (*LogNAS*), consistent with the notion that increases in consulting services may reduce the need for additional auditing services in the year of IFRS adoption, and *US_List*, capturing a firm's exposure to Sarbanes-Oxley (SOX) requirements and SEC enforcement.

Panel A, Table 2 presents the descriptive statistics for our sample of firms, including the mean and standard deviation for all variables. For binary variables, we present the count frequency. We note no particular anomalies. While mean audit fees increase over the sample period, we observe the most significant increase in the year of IFRS adoption, consistent with costly IFRS adoption. We also note a significant spike in modified opinions, from 11.7 percent of all firms in the two years prior to adoption to 15.2 percent in the year of adoption. This observation is consistent with increased uncertainty from preparers and auditors, surrounding first-time IFRS statements.

Panel B, Table 2 presents the magnitude and sign of net IFRS transmission adjustments to total equity reported by firms in their first-time IFRS statements. Of our 907 firms, 675 experienced material transitional adjustments to total equity, with 271 experiencing an increase in equity and 404 experiencing a decrease. We note that 134 firms experienced material IFRS adjustments that net to zero (i.e., net zero impact to total equity), and 98 firms reported no material adjustments whatsoever. As expected, large firms experienced the largest IFRS adjustments, while the firms that experienced no material adjustments were predominantly small, i.e., average total assets of less than \$100m. The average impact of IFRS adjustments on our sample of firms is about 2 percent of opening total assets, an economically significant amount.

To estimate the impact of IFRS adoption on audit fees, we begin with Equation (1), which includes an indicator variable, $IFRS_{it}$, that is equal to 1 if in the year of mandated IFRS adoption, and 0 otherwise. We interpret the estimated coefficient on this variable, γ_1, as the average increase in audit fees in the year of IFRS adoption relative to average fees in the pre-IFRS period. Panel A, Table 3 shows that firms incurred audit fees that are, on average, 23.7 percent higher ($\gamma_1 = 0.231$, p < 0.01) in the year of IFRS adoption.[6] Our results are similar with the inclusion of additional year dummies in the pre-IFRS period.[7]

We note that nonaudit service fees (*LogNAS*) are positive, indicating a complementary relationship between nonaudit service fees and statutory audit fees. The parameter estimates on *Loss*, *US_List*, *Opinion*, and *Big4* are all as predicted, consistent with prior research. Because the patterns in control variables are consistent across all other specifications, we focus the discussion of results on our experimental variables.

We then add three additional indicator variables capturing cross-sectional differences based on the sign of IFRS adjustments. These are reported in column (2) of Panel A, Table 3. We hand-collect data disclosed in the IFRS transition notes of first-time IFRS financial statements. This allows us to isolate those firms with material IFRS adjustments to total equity that net to zero ($Nil\_Net\_IFRS\_adj_{it}$), firms with material net positive IFRS adjustments ($Pos\_IFRS\_adj_{it}$), and firms with material net negative IFRS adjustments ($Neg\_IFRS\_adj_{it}$). The coefficients on these indicator variables, γ_2, γ_3, and γ_4, respectively, capture the impact on audit fees in the year of IFRS adoption relative to firms that experienced no material IFRS adjustments ($IFRS_{it}$). Consistent with IFRS adoption driving the observed increase in fees, we find that the coefficient γ_1 remains statistically and economically significant at 0.086 ($p < 0.01$). The coefficients on both net positive adjustments (γ_3) and negative adjustments (γ_4) were 0.156 and 0.161, respectively. These results imply an overall increase in the average level of audit fees in the year of IFRS adoption of approximately 9 percent. Firms that reported material IFRS adjustments exhibit an

6 Consistent with prior literature (for example, Craswell et al. 1995), we estimate the economic magnitude of the increase in mean audit fees associated with an indicator variable as $e^{a}-1$, where α represents the coefficient on the experimental indicator variable; in our context, the parameter estimate on *IFRS*.

7 Given that audit fees steadily increase in our pre-IFRS period, in untabulated results, we included additional year dummies to separate year t–4 and t–3. Effectively, our benchmark pre-IFRS period is then the immediate two years preceding IFRS adoption. We note that our coefficient on $IFRS_{it}$ is insignificantly different from our reported results ($\gamma_1 = 0.191$, $p < 0.01$), indicating robust results.

TABLE 2

Descriptive Statistics

Panel A: Audit Fee and Control Variables

Variable	Year *t* (IFRS Adoption) Mean	Year *t* (IFRS Adoption) Std. Dev.	*t*−1 Mean	*t*−1 Std. Dev.	*t*−2 Mean	*t*−2 Std. Dev.	*t*−3 Mean	*t*−3 Std. Dev.	*t*−4 Mean	*t*−4 Std. Dev.
Statutory Audit Fees	$337,201	$1,066,279	$274,549	$892,359	$247,789	$954,149	$216,792	$770,949	$186,854	$645,498
LogAF	11.43	1.41	11.21	1.39	11.02	1.41	10.94	1.39	10.81	1.36
Total assets	$985.3m	$5,046.9m	$831.6m	$4,311.5m	$719.5m	$3,942.6m	$777.3m	$6,121.5m	$501.2m	$2,944.3m
LogAsset	17.69	2.35	17.45	2.34	17.32	2.28	17.10	2.36	17.07	2.17
Nonaudit Service Fees	$174,735	$809,282	$179,086	$872,639	$192,212	$1,183,073	$208,673	$1,051,754	$203,174	$1,014,761
LogNAS	7.69	5.01	8.13	4.71	8.11	4.65	7.96	4.75	8.29	4.55
Quick	7.27	23.96	7.41	28.84	9.55	37.91	11.51	50.30	9.01	43.16
Accr	0.26	1.19	0.26	1.78	0.20	0.95	0.26	0.65	0.46	3.82
Rec	0.14	0.16	0.15	0.16	0.15	0.17	0.15	0.16	0.15	0.17
Inv	0.07	0.11	0.07	0.12	0.07	0.13	0.07	0.13	0.07	0.12
ROA	−0.22	1.19	−0.22	1.04	−0.20	1.15	−0.28	0.90	−0.28	0.85
Debt	0.11	0.25	0.10	0.18	0.11	0.35	0.11	0.33	0.11	0.34
Forsubs (#)	4.34	13.94	4.24	14.14	3.86	13.72	3.85	17.75	3.46	12.35
Totsubs (#)	16.00	41.81	16.26	54.27	14.79	45.00	13.11	33.03	13.08	40.98
% Loss firms	45.8%		49.8%		48.7%		56.9%		58.9%	
US_List	1.9%		1.9%		1.8%		1.8%		1.6%	
YE	84.2%		84.9%		84.4%		84.7%		85.0%	
% of modified opinions	13.7%		10.7%		10.6%		14.0%		15.0%	
% of Big 4 auditors employed	54.4%		56.5%		58.1%		59.1%		60.0%	
CPI rate (inflation)	3.81%		2.50%		2.45%		2.73%		2.84%	
No. of firms	907		907		907		907		907	

(continued on next page)

TABLE 2 (continued)

Panel B: IFRS Exposure Variables

	Total Equity (Net) Adjustments	Total Equity (Net) Increasing Adjustments	Total Equity (Net) Decreasing Adjustments	Zero (Net) Total Equity Adjustments	No IFRS Adjustments
No. of firms	675	271	404	134	98
Percent of total sample	74%	30%	44%	15%	11%
Net Adjustment to Total Equity (as percent of opening total assets)					
Mean	−2.00%	5.29%	7.97%	—	—
Std. Deviation	12.35%	10.11%	14.18%	—	—
95%	8.12%	24.94%	36.69%	—	—
Q3	0.17%	5.61%	7.99%	—	—
Median	0.00%	1.62%	2.87%	—	—
Q1	−2.19%	0.39%	0.79%	—	—
5%	−18.89%	0.02%	0.02%	—	—
Total assets (mean)	$985.3m	$1,280.1m	$1,340.1m	$93.1m	$62.1m
Total assets (std. dev.)	($5,046.9m)	($7,070.5m)	($4,826.2m)	($194.4m)	($64.8m)
% of modified opinions	13.33%	10.48%	10.67%	16.42%	27.55%
% of Big 4 auditors employed	57.54%	56.55%	67.90%	45.54%	24.48%

This table provides the mean and standard deviation statistics and count data for all variables employed in our analyses. Descriptive statistics are shown across each year employed in the analysis beginning with year *t* (year of IFRS adoption) and the preceding four years, beginning in 2001.

Variable Definitions:
LogAF = natural log of total audit fees;
LogAsset = the natural log of total assets;
LogNAS = the natural log of aggregate nonaudit service fees;
Quick = the ratio of current assets to current liabilities;
Accr = the ratio of absolute accruals to total assets;
Debt = the ratio of long-term debt to total assets;
ROA = the ratio of net income after interest and tax to total assets;
Forsubs = the number of foreign subsidiaries consolidated into the entities accounts;
Totsubs = the number of total subsidiaries consolidated into the entities accounts;
Loss = an indicator variable, equaling 1 if loss is reported in any of the prior three years, and 0 otherwise;
US_List = an indicator variable equaling 1 if the entity is listed on a U.S. stock exchange, and 0 otherwise;
YE = an indicator variable equaling 1 if the company has a June 30 year-end, and 0 otherwise;
Opinion = an indicator variable equaling 1 if a modified/qualified opinion was issued in any of the prior three years, and 0 otherwise;
Big4 = an indicator variable equaling 1 if the company employed a Big 4 auditor, and 0 otherwise; and
Adjustments = the difference in total equity measured under local GAAP and total equity measured under IFRS for the fiscal year immediately prior to mandated IFRS adoption. We collect this data from IFRS financial statements, as firms must restate comparatives in first-year IFRS-compliant financial statements.

TABLE 3

Regression Analysis (Levels)

Panel A: Level Regressions for Pooled Sample

Dependent Variable: *LogAF*	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
IFRS	0.213***	0.086**	0.169***	0.134***	
	11.22	*1.97*	*5.86*	*2.93*	
Nil_Net_IFRS_adj		0.074			
		1.01			
Pos_IFRS_adj (*dummy*)		0.156**			
		2.22			
Neg_IFRS_adj (*dummy*)		0.161**			
		2.11			
IFRS_exposure_firms			0.033***	0.074***	
			2.76	*3.55*	
IFRS ∗ *IFRS_exposure_firms*			0.019*	0.039*	
			1.68	*1.76*	
IFRS_adj_q2 (IFRS adjustments < 0.3% of opening assets)					0.073
					1.31
IFRS_adj_q3					0.144*
					1.94
IFRS_adj_q4					0.276***
					3.99
IFRS_adj_q5 (IFRS adjustments > 6.3% of opening assets)					0.211***
					2.89
LogAsset	0.298***	0.298***	0.292***	0.292***	0.292***
LogNAS	0.033***	0.033***	0.032***	0.032***	0.028***
Rec	0.948***	0.949***	0.963***	0.967***	0.817***
Inv	0.605***	0.601***	0.601***	0.594***	0.432**
Accr	0.044**	0.045**	0.044**	0.044**	0.038**
Quick	−0.002***	−0.002***	−0.002***	−0.002***	−0.004***
ROA	−0.042*	−0.042*	−0.042*	−0.043*	−0.024
Debt	0.125***	0.125***	0.119***	0.117***	0.057
Forsubs	0.232***	0.232***	0.234***	0.234***	0.197***
Totsubs	0.218***	0.216***	0.213***	0.211***	0.193***
Loss	−0.063*	−0.057*	−0.055*	−0.053	0.004
US_List	0.483***	0.489***	0.495***	0.492***	0.686***
Opinion	0.086**	0.086**	0.086**	0.086**	0.191***
Big4	0.373***	0.374***	0.371***	0.372***	0.481***
Intercept	4.607***	4.642***	4.633***	4.538***	4.847***
Year FE	Yes	Yes	Yes	Yes	NA
Adj. R^2	83.3%	83.3%	83.4%	83.4%	80.1%
n	4,535	4,535	4,535	4,535	907

(continued on next page)

additional increase of 16–17 percent. In addition, we find that 60 percent of firms experience increases in audit fees in excess of 27 percent, which is greater than the margin of error of our estimated IFRS effect. This represents the upper bound of the 95 percent confidence interval surrounding our estimate γ_1.

TABLE 3 (continued)

Panel B: Level Regressions across Different "IFRS Exposure" Subsamples

Subsample	n	IFRS Coefficient	R^2
No IFRS adjustments	490 (98 firms)	0.196*** *3.54*	59.71%
Net zero IFRS adjustments	590 (118 firms)	0.239*** *4.01*	67.21%
Largest IFRS adjustments (Q5)	850 (170 firms)	0.286*** *4.93*	83.31%

Test of difference (Largest IFRS adj − no IFRS adj) = 0.286 − 0.196 = 0.090** (z = 1.98, $p < 0.05$).

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.
This table provides coefficient estimates (t-statistics in parentheses) for the following regression model:

$$LogAF_{it} = \beta_0 + \beta_1 LogAsset_{it} + \beta_2 LogNAS_{it} + \beta_3 Rec_{it} + \beta_4 Inv_{it} + \beta_5 Accr_{it} + \beta_6 Quick_{it} + \beta_7 Debt_{it} + \beta_8 ROA_{it} + \beta_9 Forsubs_{it} + \beta_{10} Totsubs_{it} + \beta_{11} Loss_{it} + \beta_{12} US\_List_{it} + \beta_{13} Opinion_{it} + \beta_{14} Big4_{it} + \beta_{15} YE_{it} + \gamma_j IFRS\_Var_{it} + \varepsilon_{it}.$$

IFRS_Var is a given experimental variable relating to IFRS exposure. In column (1), we employ a dummy variable, *IFRS*, which is set to 1 in the year of IFRS adoption, and assigned a value of 0 otherwise. In column (2), we split the *IFRS* variable from our first specification into an additional three categories: *Pos_IFRS_adj*, which is assigned a value of 1 in the year of IFRS adoption if the firm had an equity-increasing IFRS adjustment, *Neg_IFRS_adj*, which is assigned a value of 1 in the year of adoption if the firm experienced an equity-decreasing IFRS adjustment, and *Nil_Net_IFRS_adj*, which is assigned a value of 1 in the year of IFRS adoption if the firm experienced IFRS adjustments that net to zero. We then include a firm-level variable to capture a firm's IFRS exposure based on quintile (specification 3) and tercile rankings (specification 4) of absolute IFRS adjustments. Note that nil adjustments are assigned a value of 0, and only those firms with IFRS adjustments are ranked. We also include an interaction term between our *IFRS* indicator variable and the firm-level exposure (*IFRS * IFRS_exposure_firms*), which should capture the incremental increase in mean audit fees in the year of IFRS adoption for those firms that are more highly exposed to the IFRS transition process. We expect a positive coefficient on this variable. We then turn our attention to cross-sectional variation in the level of audit fees among firms in the year of IFRS adoption. Specification (5) employs a cross-section of firms in the year of IFRS adoption (i.e., 907 firm observations), and examines the variation in audit fees based on quintiles of IFRS adjustments in the year of IFRS adoption. We include dummy variables for each quintile (excluding quintile 1, as that captures zero-adjustment firms) where *IFRS_adj_qN* is equal to 1 if the observation is in the Nth quintile of IFRS adjustments. The range of IFRS adjustments captured in each quintile is presented with variable names. We deflate audit fees and all financial variables by the CPI as of report date (base year of 2002) in order to remove the effects of inflation on our results. We report t-values in italics under parameter estimates; these are computed using clustered standard errors (firm-level). Significance levels are based on two-tailed tests.

Variable Definitions:
LogAF = the natural log of total audit fees;
LogAsset = the natural log of total assets;
LogNAS = the natural log of aggregate nonaudit service fees;
Quick = the ratio of current assets to current liabilities;
Accr = the ratio of absolute accruals to total assets;
Rec = the ratio of total receivables to total assets;
Inv = the ratio of total inventories to total assets;
Debt = the ratio of long-term debt to total assets;
ROA = the ratio of net income after interest and tax to total assets;
Forsubs = the natural log of 1 plus the number of foreign subsidiaries consolidated into the entities accounts;
Totsubs = the natural log of 1 plus the number of total subsidiaries consolidated into the entities accounts;
Loss = an indicator variable equaling 1 if a firm reported a loss in the current year;
US_List = an indicator variable equaling 1 if the entity is listed on a U.S. stock exchange and subject to SEC regulations;
Opinion = an indicator variable equaling 1 if the firm received a modified audit opinion in the current year (such as emphasis of matter, inherent uncertainty, and qualified opinions);
Big4 = an indicator variable equaling 1 if the firm employed a Big 4 auditor in the current year, and 0 otherwise; and
YE = an indicator variable equaling 1 if the company has a June 30 year-end, and 0 otherwise. We do not report parameter estimates on *YE* given its insignificance in all specifications. It is included as a control variable in all analyses.

We divide our sample according to a firm's IFRS exposure by including variable $IFRS\_Exposure_{it}$, which captures the quintile rank (column (3), Panel A, Table 3) and the tercile rank (column (4), Panel A, Table 3) of the absolute magnitude of a firm's IFRS adjustments. We then include an interaction term between $IFRS_{it}$ and the firm-level exposure ($IFRS_{it} * IFRS\_Exposure_{it}$), which captures the incremental increase in audit fees for those firms more affected by the IFRS transition process (i.e., those with larger adjustments). Consistent with our expectations, Panel A, Table 3 shows a positive and significant coefficient on the interaction term in column (3) ($\gamma_7 = 0.019$, $p < 0.05$). Thus, firms with the greatest IFRS exposure in the top quintile of IFRS adjustments incurred audit fees 7.8 percent greater than those firms with the lowest IFRS exposure in the bottom quintile. This additional fee increase is in addition to the 17 percent increase in audit fees experienced by all firms in the year of IFRS adoption. These inferences hold in column (4), where we capture IFRS exposure across high, medium, and low groups. For the above four specifications, we perform a full range of industry sensitivity testing, and our results remained unchanged in terms of sign and significance.[8] Given concerns of collinearity that may be present in these audit fee models, we review variance inflation factors (VIF) for all variables in each specification (for a brief discussion, see Lyon and Maher 2005; Judge et al. 1987). All factors are within acceptable ranges, the largest VIF of 3.21 being on *ROA*.

Our final specification in Table 3 analyzes the cross-sectional variation in the level of audit fees among firms in the year of IFRS adoption. Column (5) in Panel A, Table 3 employs the cross-section of 907 firms and examines the cross-sectional variation in audit fees based on quintiles of IFRS adjustments. We include dummy variables for each quintile, excluding quintile 1, where $IFRS\_adj\_qN_i$ is equal to 1 if the observation is in the Nth quintile of IFRS adjustments, and 0 otherwise. This specification tests for differential increases in mean audit fees across the magnitude of IFRS adjustments in the year of IFRS adoption. We find that firms with the largest IFRS adjustments, i.e., greater than 6.3 percent of assets, exhibit audit fees that are approximately 23 percent (coefficient of 0.211, $p > 0.01$) higher than firms that did not report any IFRS adjustments. In addition, we observe no reliable difference between the audit fees incurred by firms that reported zero versus very small (i.e., less than 0.3 percent of assets) IFRS adjustments. Overall, our results suggest a significant increase in audit fees in the year of IFRS adoption. Further, we find that firms with the greatest IFRS exposure exhibit the greatest increase in audit fees.

As an additional test, we partition our full sample based on the number of IFRS adjustments: no adjustments, adjustments that net to zero, and the quintile with the largest absolute IFRS adjustments. We then estimate specification (1) within in each subsample and test for differences between $IFRS_{it}$ coefficients. Panel B, Table 3 displays these results. We find cross-sectional variation consistent with IFRS adoption driving the increase in audit fees, because firms with the greatest IFRS adjustments have a significant ($z = 1.98$) increase in audit fees (11 percent higher) relative to increases experienced by firms with no material IFRS adjustments.

Changes Analysis

We next reestimate the model expressed in Equation (1) in a *changes* form to control for any time-trend and for unobservable company-specific effects. This design ensures that we are not simply capturing the impact of normal yearly increases in audit fees. Our changes empirical model is outlined below:

---

[8] Each of the ten macro GICS industries are dropped from the sample, one at a time, and the primary results are rerun, consistent with prior literature (e.g., Ferguson et al. 2003). We note that our results remain robust across all industry subsamples (IFRS coefficients ranging from 0.19 to 0.25) and significant at the 1 percent level in all specifications.

$$\begin{aligned}\Delta\_LogAF = {} & \beta_0 + \beta_1\Delta\_LogAsset_{it} + \beta_2\Delta\_LogNAS_{it} + \beta_3\Delta\_Rec_{it} + \beta_4\Delta\_Inv_{it} + \beta_5\Delta\_Quick_{it} \\ & + \beta_6\Delta\_Accr_{it} + \beta_7\Delta\_Debt_{it} + \beta_8\Delta\_ROA_{it} + \beta_9\Delta\_Forsubs_{it} + \beta_{10}\Delta\_Totsubs_{it} \\ & + \beta_{11}US\_List_{it} + \beta_{12}Loss_{it} + \beta_{13}Profit_{it} + \beta_{14}Clean_{it} + \beta_{15}Modified_{it} \\ & + \beta_{16}Big4_{it} + \beta_{17}Nonbig4_{it} + \beta_{18}YE_{it} + \beta_{19}IFRS_{it} + \varepsilon_i, \end{aligned} \quad (2)$$

where:

$\Delta\_$ = this prefix indicates that variables are measured as the difference between $t-1$ and t;
Loss = indicator variable equal to 1 if the firm reported a loss in the current year and a profit in the previous year, otherwise equal to 0;
Profit = indicator variable equal to 1 if the firm reported a profit in the current year and a loss in the previous year, otherwise equal to 0;
Clean = indicator variable equal to 1 if the firm was issued with a clean opinion in the current year and modified opinion in the previous year, otherwise equal to 0;
Modified = indicator variable equal to 1 if the firm was issued with a modified opinion in the current year and a clean opinion in the previous year, otherwise equal to 0;
Big4 = indicator variable equal to 1 if the firm switched to a Big 4 auditor in the current year from a non-Big 4, otherwise equal to 0; and
Nonbig4 = indicator variable equal to 1 if the firm switched to a non-Big 4 auditor in the current year from a Big 4, otherwise equal to 0.

All other control variables are as previously defined in the "Levels Analysis" subsection in Section IV. We include the same IFRS variables as discussed in "Levels Analysis" in order to estimate the *abnormal* increase in audit fees in the year of IFRS adoption.

Table 4 presents the results from estimating Equation (2).[9] The reported signs and significance levels on the coefficients of the control variables are in line with prior literature; notable exceptions are discussed below. In our first specification in Panel A, Table 4 we find a positive and significant coefficient on $IFRS_{it}$ of 0.078. All VIFs are within an acceptable range. Consistent with our findings in Table 3, these results find an average yearly increase in audit fees of approximately 11 percent. However, there is an *additional* 8 percent increase in the year of IFRS adoption, which we interpret as evidence of the cost of IFRS implementation incurred by firms.

In comparison to our levels analysis, we note that changes in nonaudit service fees (*LogNAS*) are positive, but no longer significant. The parameter estimates on *Loss*, *Profit*, *Clean*, and *Modified* are all consistent with prior research; for instance, a firm moving into a loss position or incurring a modified opinion will incur a sharper increase in their fees, and firms moving into a profitable position or exhibiting a clean opinion will incur a reduction in their audit fees. Another notable result from our changes analysis is the parameter estimates on control variables of Big 4 and non-Big 4. Firms that switch to Big 4 auditors from non-Big 4 auditors experienced no increase in audit fees; however, firms that switch from a Big 4 auditor to a non-Big 4 auditor experience a decrease in audit fees of approximately 48 percent (coefficient estimates ranging from -0.409 to -0.411, $p < 0.01$).[10]

---

[9] We perform a full range of industry sensitivity testing in untabulated results. Each of the ten macro GICS industries is dropped from the sample, one at a time, and the primary results rerun consistent with prior literature (e.g., Ferguson et al. 2003). We note that our results remain robust across all industry subsamples (*IFRS* coefficients ranging from 0.060 to 0.077, and significant at the 1 percent level in all specifications).

[10] We note that this result is attributable to a small portion of our sample, with only 42 instances of firms switching from non-Big 4 to Big 4 firms, and 47 instances of firms switching from Big 4 to non-Big 4 auditors. The observed decrease in audit fees resulting from a Big 4 to non-Big 4 switch is consistent with the established Big N auditor premium no longer being incurred by switching firms. Depending on the timing of the switch, this may bias down the underlying effect of the adoption of IFRS on the fees charged by these new auditors if they are engaging in price-cutting and low-balling pricing behavior. In sensitivity tests, we rerun regressions after removing the 87 instances of auditor switching within our sample period, and noted the tenor of our results remain unchanged from those obtained on the full sample.

TABLE 4

Regression Analysis (Changes)

Panel A: First Differences Regressions for Pooled Sample

Dependent Variable: ***LogAF***	(1)	(2)	(3)	(4)
IFRS	0.078***	0.033	0.061	
	3.51	*1.25*	*1.56*	
Nil_Net_IFRS_adj		0.077		
		1.67		
Pos_IFRS_adj (*dummy*)		0.082*		
		1.71		
Neg_IFRS_adj (*dummy*)		0.081*		
		1.81		
IFRS_exposure_firms			0.018**	
			2.46	
IFRS * *IFRS_exposure_firms*			0.019*	
			1.73	
IFRS_adj_q2 (IFRS adjustments < 0.3% of opening assets)				−0.004
				−0.18
IFRS_adj_q3				0.008
				1.16
IFRS_adj_q4				0.018*
				1.75
IFRS_adj_q5 (IFRS adjustments > 6.3% of opening assets)				0.031**
				1.98
Δ_LogAsset	0.188***	0.189***	0.187***	0.172***
Δ_LogNAS	0.004	0.004	0.004	0.006
Δ_Rec	−0.059	−0.062	−0.061	0.086
Δ_Inv	0.169	0.168	0.168	0.026
Δ_Accr	0.045*	0.046*	0.046*	−0.011
Δ_Quick	−0.001**	−0.001**	−0.001**	−0.002*
Δ_ROA	0.001	0.001	0.001	−0.034**
Δ_Debt	0.038	0.038	0.038	0.041
Δ_Forsubs	0.042*	0.042*	0.042*	0.037
Δ_Totsubs	0.067**	0.067***	0.067**	0.019
to_Loss	0.086**	0.088**	0.084**	0.086*
to_Profit	−0.028	−0.026	−0.031	−0.139
Clean	−0.019	−0.018	−0.015	−0.001
Modified	0.031	0.032	0.036	0.011
to_Big4	0.058	0.059	0.055	−0.055
to_Nonbig4	−0.409***	−0.411***	−0.411***	−0.335***
Intercept	0.102***	0.101***	0.063***	0.209***
Adj. R^2	14.0%	14.1%	14.2%	12.1%
n	3,628	3,628	3,628	907

(continued on next page)

Results from our remaining four specifications (columns (2) through (5), Panel A, Table 4) mirror the results of our levels analysis discussed in the "Levels Analysis" subsection in Section IV. Thus, we omit a detailed discussion of all results and draw attention to a few key findings. We note that results from column (2), Panel A, Table 4 indicate that the previously estimated *additional*

TABLE 4 (continued)

Panel B: First Differences Regressions across Different "IFRS Exposure" Subsamples

Subsample	n	IFRS Coefficient	R^2
No IFRS adjustments	392 (98 firms)	0.091** *2.47*	19.95%
Net zero IFRS adjustments	472 (118 firms)	0.114** *2.13*	19.79%
Largest IFRS adjustments (Q5)	680 (170 firms)	0.122*** *3.98*	23.10%

Test of difference (Largest IFRS adj − no IFRS adj) = 0.122 − 0.091 = 0.031*. ($z = 1.71$, $p < 0.10$).

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.
This table provides coefficient estimates (t-statistics in parentheses) for the following *changes* regression model:

$$\Delta LogAF = \beta_0 + \beta_1 \Delta LogAsset_{it} + \beta_2 \Delta LogNAS_{it} + \beta_3 \Delta Rec_{it} + \beta_4 \Delta Inv_{it} + \beta_5 \Delta Quick_{it} + \beta_6 \Delta Accr_{it} + \beta_7 \Delta Debt_{it} + \beta_8 \Delta ROA_{it} + \beta_9 \Delta Forsubs_{it} + \beta_{10} \Delta Totsubs_{it} + \beta_{11} US\_List_{it} + \beta_{12} Loss_{it} + \beta_{13} Profit_{it} + \beta_{14} Clean_{it} + \beta_{15} Modified_{it} + \beta_{16} Big4_{it} + \beta_{17} Nonbig4_{it} + \beta_{18} YE_{it} + \beta_{19} IFRS_{it} + \varepsilon_i.$$

IFRS_Var is a given experimental variable relating to IFRS exposure. In column (1), we employ a dummy variable IFRS that is set to 1 in the year of IFRS adoption, and a value of 0 otherwise. In column (2), we split IFRS from specification 1 into an additional three categories: *Pos_IFRS_adj*, which is assigned a 1 in the year of IFRS adoption if the firm had a positive IFRS adjustment to equity, *Neg_IFRS_adj*, which is assigned a value of 1 in the year of adoption if the firm experienced a negative IFRS adjustment, and *Nil_Net_IFRS_adj*, which is assigned a value of 1 in the year of IFRS adoption if the firm experienced IFRS adjustments that net to zero. We then include a firm-level variable to capture a firm's "IFRS exposure" in column (3) based on quintile rankings of absolute IFRS adjustments. Note that firms with no IFRS adjustments are assigned a value of 0. Only those firms with IFRS adjustments are ranked. We also include an interaction term between our *IFRS* indicator variable and the firm-level exposure (*IFRS * IFRS_exposure_firms*) which captures the incremental increase in mean audit fees in the year of IFRS adoption. For those firms more exposed to the IFRS transition process (i.e., those with adjustments), we expect a positive coefficient on this variable. We then turn our attention to cross-sectional variation in the level of audit fees among firms in the year of IFRS adoption. Column (4) employs a cross-section of firms in the year of IFRS adoption (i.e., 907 firm observations), and examines the cross-sectional variation in audit fees based on quintiles of IFRS adjustments in the year of IFRS adoption. We include dummy variables for each quintile (excluding quintile 1, as that captures zero adjustment firms) where *IFRS_adj_qN* is equal to 1 if the observation is in the Nth quintile of IFRS adjustments. We deflate audit fees and all financial variables by the CPI as of the report date (base year of 2002) in order to remove the effects of inflation on our results. We report t-values in italics under parameter estimates, computed using clustered standard errors (firm-level). Significance levels are based on two-tailed tests.
The Δ_ prefix indicates variables are measured as changes in value from year to year (i.e., *t*−1 to *t*).

Variable Definitions:
LogAF = the natural log of total audit fees;
LogAsset = the natural log of total assets;
LogNAS = the natural log of aggregate nonaudit service fees;
Quick = the ratio of current assets to current liabilities;
Rev = the ratio of total receivables to total assets;
Inv = the ratio of total inventories to total assets;
Debt = the ratio of long-term debt to total assets;
ROA = the ratio of net income after interest and tax to total assets;
Forsubs = the natural log of 1 plus the number of foreign subsidiaries consolidated into the entities accounts;
Totsubs = the natural log of 1 plus the number of total subsidiaries consolidated into the entities accounts;
to_Loss = an indicator variable equaling 1 if a firm reported a loss in the current year and a profit in the previous year, and 0 otherwise;
to_Profit = an indicator variable equaling 1 if a firm reported a profit in the current year and a loss in the previous year, and 0 otherwise;
Clean = an indicator variable equaling 1 if the firm received a clean audit opinion in the current year, and a modified/qualified opinion was issued in the previous year;
Modified = an indicator variable equaling 1 if the firm received a modified/qualified opinion in the current year, and a clean opinion was issued in the previous year;

(continued on next page)

TABLE 4 (continued)

Big4 = an indicator variable equaling 1 if the firm employed a Big 4 auditor, and 0 otherwise;
US_List = an indicator variable equaling 1 if the entity is listed on a U.S. stock exchange and, hence, subject to SEC regulations;
YE = an indicator variable equaling 1 if the company has a June 30 year-end, and 0 otherwise. We do not report parameter estimates on variables *YE* and *US_List*, given they are insignificant in all specifications; and
IFRS = an indicator variable defining the IFRS-adoption period, it equals 1 in the year of mandatory adoption, and 0 otherwise.

increase in fees is driven by firms that have material IFRS adjustments, as evidenced by the positive and significant coefficients on *Pos_IFRS_adj* ($\gamma_3 = 0.082$, $p > 0.10$) and *Neg_IFRS_adj* ($\gamma_4 = 0.081$, $p > 0.10$). Results from column (3) in Panel A, Table 4 reveal a positive and marginally significant coefficient on the interaction term ($\gamma_6 = 0.019$, $p < 0.10$). Consistent with the pattern in our levels analysis, this indicates firms with the greatest IFRS exposure in the top quintile of IFRS adjustments incurred greater abnormal increases in audit fees than those firms with the lowest IFRS exposure.

Overall, our results from the changes analysis are similar to those of the levels analysis, with evidence that IFRS adoption leads to significant increases in compliance costs for adopting firms.[11]

Robustness Tests

A drawback of the preceding empirical approach is that we require homogeneity of variable relationships across time (i.e., parameter estimates on the control variables are not permitted to vary between pre- and post-IFRS). When pooling pre- and post-IFRS observations, it is important to consider the underlying effect of IFRS on the measurement of empirical controls to ensure IFRS effects are appropriately estimated. This concern may be especially pertinent to the current analysis, given a number of control variables are measured using financial statement data that may not be measured consistently as a result of IFRS adjustments. We perform the following sensitivity testing: First, we perform a Chow test on the traditional levels analysis, defining the break-point as the year of IFRS adoption. We find a significant F-statistic of 9.12 ($p < 0.001$), indicating a structural break in our data that corresponds to our variable of interest, *IFRS*.

Next, we partition our sample into 15 subsamples based on the magnitude of their IFRS-exposure, defined as the absolute value of total IFRS adjustments to total equity in the year of IFRS adoption. We reestimate our column (1) specification from the levels regression, within each partition, and save the coefficients on the $IFRS_{it}$ variable. We then estimate a second-stage regression of the saved *IFRS* coefficients on changes in control variable values. We define change as the difference between mean pre-IFRS levels and post-IFRS levels, as follows:

[11] While we have documented significant increases in the average compliance costs incurred by firms in an IFRS-compliant market, our results may either be attributable to a one-off initial spike in fees, or indicate a fundamental shift upward in compliance costs going forward, given the inherent complexities and risks associated with auditing IFRS-compliant numbers. To provide some evidence on this question, we collect an additional year of data (post-adoption) for our full sample of 907 firms, and reestimate our levels and changes models (column (1)) with the inclusion of an additional indicator variable for the post-adoption year. Untabulated results indicate a marginally significant increase in the year post-IFRS adoption of an additional 11 percent (levels) and 3 percent (changes), with associated p-values of 0.021 and 0.067, respectively. This suggests that our observed increases in fees in the year of IFRS adoption has not reversed, but persisted in the post-adoption period. However, we fully acknowledge that "learning effects" and the potential exploitation of economies of scale by audit firms may not yet be revealed in the fees only one year after adoption. The post-change time-series may be too short to truly distinguish between fixed costs and transitional costs of the change (i.e., given audit cycles, one may expect a three-year transition period). A larger time-series is likely required to adequately address this question, an exercise we leave for future research.

$$IFRS\_coeff_i = \beta_0 + \beta_1 \Delta LogAsset_i + \beta_2 \Delta LogNAS_i + \beta_3 \Delta Rec_i + \beta_4 \Delta Inv_i + \beta_5 \Delta Quick_i + \beta_6 \Delta Accr_i + \beta_7 \Delta Debt_i + \beta_8 \Delta ROA_i + \beta_9 \Delta Forsubs_i + \beta_{10} \Delta Totsubs_i + \beta_{11} \Delta Loss_i + \beta_{12} \Delta YE_i + \beta_{13} \Delta Opinion_i + \beta_{14} \Delta Auditor_i + \beta_{15} \Delta US\_List_i. \quad (3)$$

Variable definitions are consistent with those reported in Table 3.

Equation (3) tests whether the variation in the observed IFRS premium is significantly influenced by any inconsistent measurement in our financial-based control variables. Table 5 shows the estimated coefficients and t-values of this second-stage regression. We find four accounting-based controls; *Rec*, *Accr*, *Quick*, and *Loss* have significant coefficients. More importantly, we find a positive and significant intercept of 0.24 ($p < 0.01$), which is not significantly different from our estimated IFRS effect from Table 3 ($\gamma_1 = 0.231$). This provides some assurance that the estimated increase in audit fees attributable to IFRS adoption is not unduly influenced by changes in the underlying measurement of control variables pre- and post-IFRS.[12]

As a final means of triangulating our observed results, we re-perform our analysis using a prediction method in which we employ parameter estimates obtained in the pre-IFRS period to compute predicted values of the mean level of audit fees in the year of IFRS adoption. Predicted values are then subtracted from observed values to obtain a measure of unexpected audit fees. In order to interpret the magnitude of the computed unexpected audit fee, we apply the methodology outlined in Van Garderen (2001) to compute an exact optimal unbiased log linear predictor. The results are consistent with those previously reported; in untabulated results, we find mean unexpected audit fees of 21.3 percent in the year of IFRS adoption.

Finally, given the possibility that some portion of the audit costs associated with IFRS adoption may be reflected in nonaudit fees, we rerun our levels and changes analysis using total fees. We find larger and marginally stronger results; our inferences remain unchanged.

V. IFRS ADOPTION COSTS AND FIRM SIZE

Professional bodies, practitioners, and members of the global financial press argue that the additional costs of preparing audited IFRS-compliant financial statements will be "heavy on smaller companies" (AICD 2005). Survey evidence from the 2005 mandatory transition to IFRS in the EU suggests transition costs of 0.31 percent of total sales for firms with less than \$700 million in sales (ICAEW 2007). In addition, small firms refer to the IFRS mandate as "overkill," citing annual reports that doubled in volume due to extensive IFRS requirements (Chong 2006).[13] These

[12] In addition, we also run a test where we estimate column (1) specification of our levels regression omitting the IFRS variable in the pre-IFRS period and post-IFRS period separately. We then take the differences in coefficients and test for significance. We find that only three of 15 control variables are significantly different in the pre- versus post-IFRS periods. Moreover, only one is based on financial statement data (*Quick*) and is weakly significant ($z = -1.83$). Overall, the relations of our control variables based on financial statement data appear consistent across pre- and post-IFRS periods. Hence, these results mitigate concerns that the potential change in measurement bases attributable to IFRS does not drive our results.

[13] While we respond to the conjecture of practitioners and executives in this section regarding the potential cost burden of IFRS adoption for small firms, we caution the reader that our results should not be viewed in isolation from other studies that document significant capital market benefits of IFRS adoption. While small firms argue they are the most affected in terms of costs, it may be that they are also significant beneficiaries of increased capital flows and the reduction in financing costs brought about by the positive effect on reporting quality. Notwithstanding, there is anecdotal evidence to suggest that may not be the case. Many small firms are not out to attract foreign investment or raise additional capital. These firms are primarily concerned with the costs and additional resources to be expended for adoption (see, for example, AICD 2005; Johnson 2009). We simply provide empirical evidence in response to these conjectures, and caution against using our results to make statements against IFRS in general or the appropriateness of IFRS adoption for smaller firms.

TABLE 5

Coefficient Estimates for the Audit Fee Levels Regression Model and Chow Test

Intercept	Δ_LogAsset	Δ_LogNAS	Δ_Rec	Δ_Inv	Δ_Quick	Δ_Accr	Δ_ROA	Δ_Debt	Δ_Forsubs	Δ_Totsubs	US_List	Loss	YE	Opinion	Auditor
0.23**	0.010	0.010	0.021**	0.019	0.0329*	0.011*	−0.005	−0.002	0.022	−0.007	0.002	0.011*	0.021	−0.011	−0.012
(19.78)	(0.30)	(1.26)	(2.01)	(1.61)	(1.91)	(1.81)	(−0.91)	(−0.69)	(0.35)	(1.41)	(1.29)	(1.76)	(1.16)	(−1.59)	(−1.44)

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.
This table provides coefficient estimates (t-statistics in parentheses) for the following regression model:

$$IFRS\_coeff_i = \beta_0 + \beta_1\Delta\_LogAsset_i + \beta_2\Delta\_LogNAS_i + \beta_3\Delta\_Rec_i + \beta_4\Delta\_Inv_i + \beta_5\Delta\_Quick_i + \beta_6\Delta\_Accr_i + \beta_7\Delta\_Debt_i + \beta_8\Delta\_ROA_i + \beta_9\Delta\_Forsubs_i + \beta_{10}\Delta\_Totsubs_i + \beta_{11}\Delta\_Loss_i + \beta_{12}\Delta\_YE_i + \beta_{13}\Delta\_Opinion_i + \beta_{14}\Delta\_Auditor_i + \beta_{15}\Delta\_US\_List_i. \quad (3)$$

We report t-values in parentheses under parameter estimates; they are computed using clustered standard errors (company and year-level).
We note that the adjusted R^2 of the regression estimation is 9.1 percent, estimated over 907 observations.

Variable Definitions:
The prefix Δ_ indicates variables are measured as changes in value from pre-IFRS period (mean value from $t-3$ through $t-1$) to the IFRS-period ($t = 0$);
LogAsset = the natural log of total assets;
LogNAS = the natural log of aggregate nonaudit service fees;
Rec = the ratio of total receivables to total assets;
Inv = the ratio of total inventory to total assets;
Quick = the ratio of current assets to current liabilities;
Accr = the ratio of absolute accruals to total assets;
Debt = the ratio of long-term debt to total assets;
ROA = the ratio of net income after interest and tax to total assets;
Forsubs = the natural log of 1 plus the number of foreign subsidiaries consolidated into the entities accounts (we report the mean and standard deviation of the actual number of foreign subsidiaries in parentheses);
Totsubs = the natural log of 1 plus the number of total subsidiaries consolidated into the entities accounts (we report the mean and standard deviation of the actual number of total subsidiaries in parentheses);
US_List = an indicator variable equaling 1 if the entity is listed on a US stock exchange and, hence, subject to SEC regulations;
Loss = an indicator variable equaling 1 if a firm reported a loss in the current year, and 0 otherwise;
Opinion = an indicator variable equaling 1 if the firm received a modified/qualified opinion in the current year, and 0 otherwise;
Auditor = an indicator variable equaling 1 if the firm employs a Big 4 auditor in the current year, and 0 otherwise;
YE = an indicator variable equaling 1 if the company has a June 30 year-end, and 0 otherwise;
IFRS = an indicator variable defining the IFRS adoption period. It equals 1 in the year of mandatory adoption, and 0 otherwise; and
Coeff_IFRS = saved coefficients on *IFRS* from a first-stage set of regressions of Equation (3), the "levels" audit fee model runs within 15 portfolios sorted on the absolute value of IFRS adjustments made to total equity.

sentiments are echoed in the U.S., with a recent Accenture survey finding that U.S. executives actually expect to pay *more* than their EU counterparts, estimating IFRS adoption costs of up to 0.7 percent of annual revenues depending on firm size (Johnson 2009).

On the other hand, standard setters have asserted that small companies are in for a pleasant surprise, as "[IFRS] is not such an onerous thing as they were first led to believe" (Andrews 2005, 86). In response to the international debate, the International Accounting Standards Board (IASB) launched a project to develop specific IFRS for Small and Medium-sized Entities (SMEs), with the goal to reduce the preparation and compliance costs associated with full-scale IFRS adoption. However, there has been little evidence to support conjectures of either side. We provide empirical evidence on this issue. Partitioning the sample based on size, we examine the differential impact of IFRS on the audit costs incurred by small versus large firms. Panel A, Table 6 describes our size partitions based on the distribution of firms' total assets in the year preceding IFRS adoption (i.e., firm size measured under local GAAP).

Panel B, Table 6 presents results for our levels analysis across size partitions. Panel C, Table 6 shows our results for our changes analysis. We find the firms in the smallest size quintile exhibit the largest increase in audit fees in the year of IFRS adoption ($\gamma_1 = 0.266$, $p < 0.01$) relative to firms in the largest quintile ($\gamma_1 = 0.181$, $p < 0.01$). Our results find that small firms incur an average increase in audit fees of approximately 30 percent relative to the pre-IFRS period, while large firms incurred an average increase in audit costs of approximately 19.8 percent. Additional analysis on firm size reveals a monotonic pattern in the observed magnitude of the IFRS costs. We classify firms with assets less than \$100m as small, firms with assets between \$100m and \$500m as medium, and those with assets above \$500m as large. Overall, we find that the greatest increases in audit fees are borne by small firms. Further, breaking down our IFRS indicator variable to capture the differential impact on firms based on IFRS adjustments, we find that small firms with no IFRS adjustments still experience a significant increase in their audit fees of approximately 10 percent (coefficient on *No_IFRS_adj* = 0.095, $p < 0.10$) in the year of adoption. Whereas small firms with material IFRS adjustments experience an average increase in audit fees of approximately 36 percent in the year of adoption. Conversely, large firms with the least exposure to IFRS (i.e., net effect of all IFRS adjustments on total equity of zero) actually incurred a significant decrease in audit fees. Large firms only incur a significant increase in audit fees if they report material and negative adjustments, with a significant coefficient on *Neg_IFRS_adj* ranging between 0.292 and 0.337. In untabulated results, we repeat the size analysis controlling for potential industry-specific effects, noting qualitatively similar results to those reported in Table 5.

Panel C, Table 6 provides the results from our size partitions across our changes analysis. We find similar results to those reported in Panel B. Specifically, we lose significance on the *IFRS* variable in our column (1) specification for firms in the top size quintile. But we find significant additional increases in audit fees for small firms in the smallest size quintile (coefficient estimate of 0.14, $p < 0.01$). Consistent with results in Panel B, we find the small firms incur an abnormal increase in audit fees in the year of adoption of approximately 16 percent relative to expected yearly increases. No abnormal increase is detectable for the large firms. Large firms only incur abnormal increases in audit fees when they report material IFRS adjustments, an additional 9 percent increase in audit fees above expected yearly increases. Interestingly, small firms with no material IFRS adjustments still incur an abnormal increase of 10 percent in the year of adoption (coefficient on *No_IFRS_adj* of 0.095, $p < 0.10$), while small firms with material adjustments incur the most significant jump in audit fees in the year of adoption, increasing by an additional 21 percent. Overall, these results provide evidence consistent with the conjecture that small firms incur disproportionately more IFRS costs than large firms.

TABLE 6

IFRS Related Fee Changes for Small versus Large Firms

Panel A: Details of Size Partitions

Size Classification	Q1 (Smallest)	Q5 (Largest)	Assets < $100m	Asset $100m–500m	Assets > $500m
Avg. Total assets	$14.61m	$2,660.12m	$23.42m	$214.08m	$5,249.26m
Abs_IFRS_adj	0.71%	1.68%	1.23%	1.55%	1.67%
No. of firms	181	180	602	180	125

Panel B: Levels Analysis across Small, Medium, and Large Firms

	Q1 (Small Firms)		Q5 (Large Firms)		Assets < $100m		Asset $100m–$500m		Assets > $500m	
IFRS Variables:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
IFRS	0.266***	—	0.181***	—	0.246***	—	0.201***	—	0.172**	—
	6.08		*4.64*		*10.11*		*4.34*		*4.98*	
No_IFRS_adj	—	0.154*	—	NA	—	0.095*	—	−0.153***	—	NA
		2.31				*1.75*		*−3.47*		
Nil_Net_IFRS_adj	—	0.171*	—	−0.072***	—	0.162***	—	0.093	—	−0.153**
		1.74		*−3.14*		*2.75*		*0.39*		*−2.31*
Pos_IFRS_adj (dummy)	—	0.513***	—	0.071	—	0.341***	—	0.166*	—	0.041
		5.56		*0.46*		*7.65*		*1.85*		*0.15*
Neg_IFRS_adj (dummy)	—	0.351***	—	0.292***	—	0.229***	—	0.262**	—	0.337***
		3.86		*5.42*		*5.74*		*3.54*		*6.13*
Controls included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	51.7%	51.4%	79.5%	79.8%	65.5%	65.6%	69.1%	69.1%	79.4%	79.7%
n	905	905	900	900	3,010	3,010	900	900	625	625

(continued on next page)

TABLE 6 (continued)

Panel C: Changes Analysis across Small, Medium, and Large Firms

	Q1 (Small Firms)		Q5 (Large Firms)		Assets < $100m		Asset $100m–$500m		Assets > $500m	
IFRS Variables:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
IFRS	0.141***	—	0.181***	—	0.246***	—	0.201***	—	0.172***	—
	3.25		*4.64*		*10.11*		*4.34*		*4.98*	
No_IFRS_adj	—	0.096*	—	N/A	—	0.095*	—	−0.153***	—	N/A
		1.86				*1.75*		*−3.47*		
Nil_Net_IFRS_adj	—	0.096	—	−0.072***	—	0.162***	—	0.093	—	−0.153**
		1.66		*−3.14*		*2.75*		*0.39*		*−2.31*
Pos_IFRS_adj (dummy)	—	0.208***	—	0.071	—	0.341***	—	0.166*	—	0.041
		2.96		*0.46*		*7.65*		*1.85*		*0.15*
Neg_IFRS_adj (dummy)	—	0.198***	—	0.292***	—	0.229***	—	0.262**	—	0.337***
		2.89		*5.42*		*5.74*		*3.54*		*6.13*
Controls included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	18.5%	18.8%	79.5%	79.8%	65.5%	65.6%	69.1%	69.1%	79.4%	79.7%
n	724	724	720	720	2,408	2,408	720	720	500	50

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.

This table provides coefficient estimates for the levels models (Equation (1)) as specified in Table 3 and changes models (Equation (2)) as specified in Table 4, across partitions based on firm size. We partition firms into balanced size quintiles based on the total assets in the year prior to IFRS transition. In Panel A, we provide details of our size partitions, and report the average absolute IFRS adjustment to total equity reported by firms. We report our levels analysis in Panel B and our changes analysis in Panel C. We employ a number of different experimental variables to capture the increase in audit fees attributable to IFRS as per Table 3 and Table 4. We define these as follows: a simple dummy variable, *IFRS*, which is set to 1 in the year of IFRS adoption, and assigned a value of 0 otherwise. We then partition our *IFRS* variable into four distinct dummy variables: *No_IFRS_adj*, which is assigned a value of 1 if the firm reported no IFRS adjustments, and 0 otherwise; *Nil_Net_IFRS_adj*, which is assigned a value of 1 in the year of IFRS adoption if the firm experienced IFRS adjustments that net to zero; *Pos_IFRS_adj*, which is assigned a value of 1 in the year of IFRS adoption if the firm had an equity-increasing IFRS adjustment; and *Neg_IFRS_adj*, which is assigned a value of 1 in the year of adoption if the firm experienced an equity-decreasing IFRS adjustment. Control variables employed in Panels B and C are those used in Table 3 and Table 4, respectively. We deflate audit fees and all financial variables by the CPI as of the report date (base year of 2002) in order to remove the effects of inflation on our results. We report t-values in italics under parameter estimates, computed using clustered standard errors (firm-level).

VI. THE IMPACT OF SPECIFIC IFRS STANDARDS

Finally, we provide insight concerning the aspects of IFRS that are the most costly. Prior professional and academic literature provides an initial understanding as to which specific IFRS had the most impact. Reports issued by professional bodies and academics (e.g., Ernst & Young 2005; Jubb 2005) consistently identify six key accounting policies of IFRS that have the most material impact on financial statement preparation.[14] These standards relate to *Share-Based Payments* (IFRS 2/AASB 2), *Income Taxes* (IAS 12/AASB 112), *Employee Benefits* (IAS 19/AASB 119), *Impairment* (IAS 36/AASB 136), *Intangibles* (IAS 38/AASB 138), and *Financial Instruments* (IAS 32/39; AASB 132/139).[15]

The above list is generated from empirical analysis of financial statements and survey responses of firm executives; however, the examination of audit costs is reliant on auditor behavior. To this end, we broaden the existing literature by surveying 45 professional auditors from a Big 4 accounting firm in order to assess the direct effect of IFRS on the audit function. These auditors were asked to rate the incremental audit effort/risk associated with the previously identified six standards. Our participants were asked to rate, from 1 through 10, and rank each IFRS according to the additional effort required to satisfy the new requirements *relative* to previous Australian GAAP. This allows us to capture the incremental effects of IFRS adoption on the compliance procedures undertaken by auditors and to isolate the specific IFRS requirements that are likely to result in the greatest costs to firms. Appendix A provides a brief discussion on the significant changes embodied within each of the six IFRS discussed here. The complete research instrument is available from the authors upon request.

Table 7 shows a statistical summary of the responses from our survey questionnaires. Of the 45 professional accountants surveyed, we received 31 usable responses consisting of: 12 Assistant Managers, averaging 3.75 years of experience, seven Managers with approximately six years of experience, seven Senior Managers with ten years of experience, and five Partners with approximately 18 years of experience. All respondents are members of professional accounting bodies and have attended in-house IFRS training. Panel B, Table 7 shows that the standard governing the recognition and measurement of financial instruments (AASB 139) attracts the highest rating of 8.3 out of 10. Coupled with a mean rating of 7.3 on AASB 132 (AASB 2004e), the standard governing the presentation of financial instruments, these results demonstrate that auditors expend the greatest level of audit effort in complying with the new IFRS requirements for financial instruments. Moreover, 68 percent of respondents ranked AASB 132 and AASB 139 (AASB 2004f) as having the greatest impact on audit effort (Panel C, Table 7). Panel B also shows that the requirements in AASB 2 (AASB 2004a) encompassing share-based incentives as having a

---

[14] A number of studies undertaken by the accounting profession provide insight into the potential impacts of IFRS adoption on financial reporting. A survey commissioned by the Institute of Chartered Accounting of Australia (ICAA) of over 200 Australian businesses raised concerns about the onerous IFRS requirements surrounding the application of income taxes, financial instruments, and impairment. The TCG Report released by Ernst & Young in 2005 reports similar findings, suggesting that share-based incentives, income taxes, defined benefit plans, and financial instruments have the most significant impact on the financial statement disclosures of proposed IFRS adjustments of Australian companies (Ernst and Young 2005). Jubb (2005) provides similar evidence based on her empirical investigations of the anticipated IFRS adjustments disclosed by 808 companies in their "expected impacts of transition" disclosures, required by AASB 1047 in the year leading to full adoption. Her results consistently identified five key accounting policy differences arising from transition to IFRS, namely: income tax, asset impairment, share-based remuneration, financial instruments, and goodwill on consolidation.

[15] We use the references to the IFRS and IAS standards interchangeably with the Australian equivalents to IFRS (AASBs). The specific standards examined in this study are nearly identical with the IAS/IFRS equivalent, with minor variations in AASB 112 (AASB 2004c) given the differing tax legislation within Australia. For detailed commentary between IFRS as released by the IASB and Australian equivalents to IFRS, see publications issued by the Big 4 at: http://www.iasplus.com/en/binary/au/0509differences.pdf/view

TABLE 7

Summary of Survey Opinions

Panel A: Biographical Information of Respondents

	Count Number of Responses	Percent	Mean Years of Experience	Mean Number of IFRS Training
Assistant Managers	12	38	3.75	4
Managers	7	23	6.07	4–5
Senior Managers	7	23	10.10	>5
Partners	5	16	18.00	>5
	31	100		

Panel B: Mean (Standard Deviations) of Audit Effort/Complexity Ratings Assigned by Respondents

	Full Sample	Lower-Level Staff^{+}	Senior-Level Staff^{-}	t-statistic
Number of Observations	22	6	16	
AASB 2	7.7	6.3	8.3	2.02**
"Share-Based Payments"	(1.6)	(2.2)	(1.0)	
Number of Observations	31	12	19	
AASB 119	4.9	3.4	5.8	3.01**
"Employee Benefits"	(2.4)	(2.2)	(2.1)	
AASB 112	6.5	6.2	6.6	0.76
"Income Taxation"	(1.7)	(1.6)	(1.7)	
AASB 132	7.3	7.1	7.5	0.54
"Financial Instruments—Disclosure"	(1.9)	(2.3)	(1.6)	
AASB 139	8.3	7.2	9.1	2.94**
"Financial Instruments—Recognition and Measurement"	(1.7)	(1.8)	(1.1)	
AASB 136	6.8	6.2	7.2	1.53*
"Impairments"	(1.6)	(1.8)	(1.5)	
AASB 138	6.3	5.8	6.7	1.30
"Intangibles"	(2.0)	(2.0)	(1.8)	

Panel C: Top Three Rankings of New IFRS with Regard to Audit Effort

	Number of Experts					
	Primary (1)	Percent	Secondary (2)	Percent	Tertiary (3)	Percent
AASB 2 (n = 22)	5	23%	5	23%	7	32%
AASB 119 (n = 31)	0	0%	3	10%	1	3%
AASB 112 (n = 31)	1	3%	4	13%	4	13%
AASB 132 (n = 31)	4	13%	9	29%	5	16%
AASB 139 (n = 31)	17	55%	6	20%	4	13%
AASB 136 (n = 31)	2	6%	2	6%	5	16%
AASB 138 (n = 31)	2	6%	2	6%	5	16%

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.

Panel A reports biographical information of expert respondents, such as their years of audit experience and number of IFRS training sessions attended, based on their level within the organization.

(continued on next page)

TABLE 7 (continued)

Panel B reports the mean and standard deviation of expert ratings of the incremental audit effort/risk associated with new IFRS requirements within the relevant accounting standards: AASB 2, AASB 119, AASB 112, AASB 132, AASB 139, AASB 136, and AASB 138. The rating ranges from 0 to 10, where 0 indicates no perceived impact on audit effort/complexity, and 10 indicates an extremely high perceived impact on audit effort/complexity. We further report the mean ratings of lower-level staff (Assistant Managers with an average of 3.75 years of audit experience) and mean ratings of higher-level staff (Managers and above, with an average of 11 years of audit experience). The t-statistics reported represent the statistical difference between the ratings assigned by lower-level audit staff and upper-level audit staff. We find that an expert's level of experience plays a significant role in how IFRS impacts audit engagements.

Panel C reports the frequency (percentages) of respondents that ranked relevant standards as: (1) primary impact, indicating the standard as having the most significant impact on auditors; (2) secondary impact, indicating the standard as having the second most significant impact on auditors; and (3) tertiary impact, indicating the standard as having the third most significant impact on auditors.

Numbers in parentheses represent the number of experts that ranked each standard. Note that AASB 2 has only been ranked by 22 of the 31 experts. This is due to nine respondents placing an NA within the box. A review of respondents' comments revealed that these auditors had not yet had direct experience with applying this standard in an audit engagement and, therefore, were not in a position to rate it.

significant effect on audit effort, second behind financial instruments, with a mean rating of 7.7. Twenty-three percent of experts rank AASB 2 as the most significant impact on audit effort and required expertise. Impairment also had a significant impact on audit effort, with AASB 136 (AASB 2004d) given a mean rating of 6.8.

Based on the above insights, we reestimate our traditional levels regression model as outlined in Equation (1), and include an additional variable capturing variation across firms' exposure to IFRS standards expected to have the greatest impact, *IFRS_Score*. Specifically, this variable is an *aggregate* measure of the extent a firm is subject to IFRS provisions, defined below. We include an interaction term in the year of IFRS adoption to isolate the impact in the year of transition, as follows:

$$\begin{aligned} LogAF_{it} = {} & \beta_0 + \beta_1 LogAsset_{it} + \beta_2 LogNAS_{it} + \beta_3 Quick_{it} + \beta_4 Accr_{it} + \beta_5 Rec_{it} + \beta_6 Inv_{it} \\ & + \beta_7 Debt_{it} + \beta_8 ROA_{it} + \beta_9 Forsubs_{it} + \beta_{10} Totsubs_{it} + \beta_{11} US\_List_{it} \\ & + \beta_{12} Loss_{it} + \beta_{13} Opinion_{it} + \beta_{14} Big4_{it} + \beta_{15} YE_{it} + \gamma_1 IFRS_{it} + \varphi_1 IFRS\_Score_{it} \\ & + \varphi_2 IFRS{*}IFRS\_Score_{it} + \varepsilon_{it}, \end{aligned} \tag{4}$$

where:

IFRS_Score = sum of *Hedge*, *R_Fin*, *Fin_inst*, *SBI*, *Tax*, *Gwill*, and *Intan*. This can take a maximum value of 11, being highest exposure, to a minimum value of 0, no exposure; where:

Hedge = indicator variable equal to 1 if the company applied hedge accounting under AASB 139, otherwise equals 0;

R_Fin = quintile rank of gross financial assets and liabilities (scaled by opening total assets);[16]

[16] To measure the impact of AASB 132/139 on audit fees, we used three approaches: (1) quintile rank of gross financial assets and liabilities, in the year of IFRS adoption; (2) whether the company reports financial instrument derivatives on the balance sheet; and (3) whether the company applied hedge accounting (*Hedge*). We posit that measuring the gross value of financial assets and liabilities captures the increased risk and audit effort required, as IFRS requires significant disclosures and fair value measurement. More complex measurements are associated with financial instrument derivatives; therefore, derivatives represent greater auditor effort and risk. The requirements governing hedge accounting encompass significant changes from previous local GAAP; namely, more stringent requirements over hedge effectiveness and designation of hedged relationships. If a company is applying hedge accounting as governed by AASB 132/139, they are required to disclose this information in their accounting policies described in the annual report, allowing us to ascertain which companies are undertaking hedge accounting.

TABLE 8

Characteristics of "High Impact" IFRS Firms

Panel A: Descriptive Statistics for Subsets of High "*IFRS_Score*" Firms

Variable	Hedge Mean	Hedge Std. Dev.	Derivative Financial Instruments Mean	Derivative Financial Instruments Std. Dev.	Share-Based Payments Mean	Share-Based Payments Std. Dev.	Tax Mean	Tax Std. Dev.	Goodwill Mean	Goodwill Std. Dev.	Intangibles Mean	Intangibles Std. Dev.
Audit Fees ('000)	$527	$1,167	$897	$1,652	$691	$1,399	$634	$1,100	$701	$1,245	$699	$1,458
Total Assets	$2,129m	$9,380m	$3,132m	$9,400m	$2,111m	$8,756m	$1,950m	$6,931m	$2,152m	$7,888m	$1,978m	$7,555m
LogNAS	6.23	4.56	7.12	4.67	8.56	4.33	7.68	5.65	9.29	4.35	8.23	4.67
Quick	9.23	31.43	6.75	29.86	7.35	28.40	6.79	37.56	8.25	34.67	7.98	41.23
Accr	0.41	5.67	0.51	6.58	0.45	7.11	0.49	5.31	0.55	7.86	0.49	6.82
Debt	0.12	0.31	0.11	0.33	0.11	0.29	0.11	0.37	0.09	0.39	0.11	0.23
ROA	−0.69	11.34	−0.55	9.24	−0.45	6.48	−0.72	9.18	−0.51	6.33	−0.62	15.33

Panel B: Frequency of Firms Subject to High-Impact Requirements, and Distribution of *IFRS_Score* across Size Quintiles

	Hedge	*Rank_Fin_inst*	*Deriv_Fin_inst*	*SBI*	*Tax*	*Gwill*	*Intan*	*IFRS_Score* (Mean)
Q1 (Smallest firms)	82	2.74	97	89	68	94	76	6.01
Q2	73	3.22	80	91	74	97	85	7.21
Q3	63	2.87	75	142	60	89	70	7.01
Q4	101	3.22	123	113	88	111	95	8.75
Q5 (Largest firms)	111	4.3	132	141	95	121	100	9.88

This table provides summary statistics (mean, median, standard deviation, and counts) for the subsets of firms that are identified as having been exposed to high-impact areas of IFRS/AIFRS employed in the estimation of Equation (2).

The key high-exposure determinants are: *Hedge*, indicating firms applied Hedge accounting under AASB 139 (IAS 39); *Rank_Fin_inst*, quintile rank of gross financial assets and liabilities (scaled by opening total assets); *Deriv_Fin_inst*, indicating firms that are subject to AASB 132/139 (IAS 32/39); *SBI*, indicating a subset of firms that recognize an IFRS adjustment pertaining to share-based payment transactions due to IFRS requirements under AASB 2 (IFRS 2); *Tax*, indicating the subset of firms that recorded IFRS adjustments pertaining to taxes under AASB 112 (IAS 12); *Gwill*, indicating the subset of firms that recognized IFRS adjustments over goodwill balances; *Intan*, indicating the subset of firms that recognized intangible balances (excluding Goodwill) in the period prior to IFRS adoption and, hence, are subject to AASB 138 (IAS 38); and *IFRS_Score* is defined as the sum of: *Hedge, R_Fin, Fin_inst, SBI, Tax, Gwill, and Intan.* This can take a maximum value of 11 (high exposure) and a minimum value of 0 (lowest exposure).

Fin_inst = indicator variable equal to 1 if the company recognizes derivative financial instruments in the year of IFRS adoption, otherwise equals 0;
SBI = indicator variable equal to 1 if the company applies AASB 2, otherwise equals 0;[17]
Tax = indicator variable equal to 1 if the company recorded IFRS adjustments in relation to AASB 112 (AASB 2004b), otherwise equal to 0;[18]
Gwill = indicator variable equal to 1 if the company recorded goodwill balance, otherwise equal to 0;[19] and
Intan = indicator variable equal to 1 if the company recognized intangible assets in prior year under previous AGAAP, otherwise equal to 0.[20]

We expect the coefficient on φ_2 to be significantly positive, indicating that companies exposed to these standards incur a greater increase in their audit fees in the year of IFRS transition. Table 8 presents descriptive statistics for firms subject to the identified *high-impact* IFRS requirements, along with the number of firms subject to these requirements. We also report descriptive statistics across size quintiles. Notably, a significant number of small firms are subject to *high-impact* requirements, as are the majority of large firms. While it may be reasonable to assume small firms are less likely to undertake complex transactions, such as financial instruments or acquisitions giving rise to Goodwill, it appears the distribution of firms subject to *high-impact* IFRSs is fairly consistent across size quintiles. This alleviates any potential concerns of low power within our quintile comparisons. However, we do see a significant decrease in the average *IFRS_Score* across size quintiles with the smallest firms, i.e., those in the bottom quintile, exhibiting an average *IFRS_Score* of 6.01, as compared to 9.88 for firms in the top quintile.

Panel A, Table 9 presents the coefficient estimates on the *IFRS* term *IFRS_Score* and the interaction term. We omit reporting estimates on the control variables for ease of exposition. All variables within this analysis exhibit VIF scores well within the acceptable range. The coefficient on *IFRS* is positive and significant (coefficient estimate of 0.15, $p < 0.01$), and is somewhat attenuated from results reported in Table 3: reduced from 23.7 percent to 16.1 percent. The coefficient on the *IFRS_Score* variable is positive and significant (coefficient estimate of 0.16, $p < 0.01$), indicating that high *IFRS_Score* firms exhibit higher audit fees in general across our sample period. Interestingly, this indicates that *IFRS_Score* is capturing an aspect of perceived risk/complexity that auditors price above and beyond the traditional determinants that are based primarily on reported financial statement amounts. In line with our predictions, the interaction effect is positive and significant (coefficient estimate of 0.08, $p < 0.05$), indicating that the greater a firm's exposure to these six IFRSs, the greater the costs incurred by firms in the year of adoption.

---

[17] With respect to AASB 2, we reviewed the IFRS transition notes to the financial statements to identify which companies recognized adjustments related to share-based payments. Where a company recognized adjustments in relation to AASB 2, we assigned it a value of 1, and 0 in all other circumstances.

[18] In measuring the impact of AASB 112, we reviewed the mandatory disclosures within the notes to the financial statements that detail all material changes to both profit and equity arising from the transition to IFRS. Companies who reported such material adjustments were assigned a value of 1 (*Tax*), while all other companies were assigned a value of 0.

[19] The new IFRS on impairment applies primarily to indefinite life intangibles (i.e., goodwill balances). The risk posed by this standard relates primarily to impairment tests of goodwill balances. Therefore, if a company reports a goodwill balance in the year of IFRS transition, it was assigned a value of 1 (*Gwill*), and 0 in all other circumstances.

[20] AASB 138 (AASB 2004g) requires the derecognizing of self-generated intangible assets and any previous capitalized Research and Development expenditure that does not meet IFRS requirements. To capture all companies affected by the AASB 138, we reviewed annual reports in the current and preceding year of IFRS adoption to identify which companies held intangible assets other than goodwill. Such companies were assigned a value of 1 (*Intan*), and 0 in all other circumstances.

TABLE 9

IFRS Score and the Costs of IFRS Adoption

Panel A: Coefficient Estimates for All Data

	IFRS	*IFRS_Score*	*Interaction*	Adj. R^2
n = 4,535 (907 firms)	0.15***	0.16**	0.08**	84.12%
	(4.11)	(2.85)	(2.18)	

Panel B: Coefficient Estimates for Size Quintiles

		IFRS	*IFRS_Score*	*Interaction*	Adj. R^2
Q1 (Smallest)	n =905	0.22***	0.10*	0.05	58.11%
	(181 firms)	(4.57)	(1.71)	(1.56)	
Q2	n =910	0.19**	0.18*	0.07	68.22%
	(182 firms)	(3.11)	(1.84)	(1.43)	
Q3	n =910	0.16*	0.13*	0.05*	71.48%
	(182 firms)	(1.91)	(1.81)	(1.78)	
Q4	n = 910	0.14**	0.15**	0.07*	79.83%
	(182 firms)	(2.09)	(1.97)	(1.91)	
Q5 (Largest)	n =900	0.11	0.17***	0.09***	80.15%
	(180 firms)	(1.69)	(3.87)	(3.11)	

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.
This table provides coefficient estimates (t-statistics in parentheses) for the following regression model:

$$\begin{aligned} LogAF_{it} = {} & \beta_0 + \beta_1 LogAsset_{it} + \beta_2 LogNAS_{it} + \beta_3 Rec_{it} + \beta_4 Inv_{it} + \beta_5 Accr_{it} + \beta_6 Quick_{it} + \beta_7 Debt_{it} + \beta_8 ROA_{it} \\ & + \beta_9 Forsubs_{it} + \beta_{10} Totsubs_{it} + \beta_{11} Loss_{it} + \beta_{12} US\_List_{it} + \beta_{13} Opinion_{it} + \beta_{14} Big4_{it} + \beta_{15} YE_{it} \\ & + \beta_{16} IFRS_{it} + \beta_{17} IFRS\_Exposure_{it} + \beta_{18} IFRS{*}IFRS\_Exposure_{it} + \varepsilon_{it}. \end{aligned} \tag{4}$$

The control variables are consistent with control variables throughout this paper.
Panel A shows the results on *IFRS*, *IFRS_Score*, and the interaction term of IFRS year and IFRS exposure. We have not reported the estimates on the control variables in the interest of parsimony; however, the results on control variables appear quantitatively similar with results in Table 3 and Table 6.
Panel B shows coefficient estimates based on size quintiles as per Table 5.
We report t-values based on clustered standard errors (firm-level).
Control variables are as defined in Table 3.

Variable Definitions:
IFRS_Score = the sum of *Hedge, R_Fin, Fin_inst, SBI, Tax, Gwill, and Intan*. This can take a maximum value of 11 (high exposure) and minimum value of 0 (low exposure);
Hedge = an indicator variable equal to 1 if the company applied hedge accounting under AASB 139, otherwise equals 0;
R_Fin = a quintile rank of gross financial assets and liabilities (scaled by opening total assets);
Fin_inst = an indicator variable equal to 1 if the company recognizes derivative financial instruments in the year of IFRS adoption, otherwise equals 0;
SBI = an indicator variable equal to 1 if the company applies AASB 2, otherwise equal to 0;
Tax = an indicator variable equal to 1 if the company recorded IFRS adjustments in relation to AASB 112, otherwise equal to 0;
Gwill = an indicator variable equal to 1 if the company recorded goodwill balance, otherwise equal to 0;
Intan = an indicator variable equal to 1 if the company recognized intangible assets in prior year (i.e., under previous GAAP), otherwise equal to 0; and
IFRS = an indicator variable that equals 1 in the year of mandatory IFRS adoption, and 0 otherwise. The interaction term *IFRS * IFRS_Score* captures the incremental increase in audit fees experienced by high-exposure firms. We predict a positive and significant coefficient for this interaction term.

Panel B, Table 9 presents the results across size quintiles. The magnitude and significance on the *IFRS_Score* variable drops 0.10 and is now only marginally significant, with $p < 0.10$ for firms in the smallest size quintile. However, the *IFRS* variable increases and remains significantly positive (coefficient estimate of 0.22, $p < 0.01$). This indicates that small firms with a high *IFRS_Score* do not exhibit significant increases in audit fees, in contrast to small firms with a low *IFRS_Score*. Reconciling this finding with our main results, it appears that small firms exhibit higher increases in audit fees in the year of adoption, independent to their level of exposure to the more complex IFRSs. A potential explanation for this result may be a *fixed* IFRS switching cost. That is, there is likely an increase in the audit resources necessary to fulfill IFRS requirements and, hence, a fixed component to IFRS costs. We note, however, that large firms do not appear to exhibit a fixed IFRS cost. The coefficient estimate on the *IFRS* variable is not significant for firms in the largest size quintile. For firms in the two largest size quintiles, we find that firms' exposure to complex IFRS, *IFRS_Score*, drives the observed post-IFRS increase in audit fees, as evidenced by the positive and significant interaction term. Our results are also consistent with allegations appearing in the financial press of opportunistic pricing by audit firms directed at small firms; an interesting area for future research.

Our empirical results confirm our survey responses, and find that increases in audit costs around IFRS adoption can be traced primarily to requirements regarding share-based payments, hedge accounting designations, financial instruments, goodwill and intangible balances, and IFRS tax adjustments.

VII. SUMMARY AND CONCLUSIONS

This paper provides initial evidence on the costs associated with mandated IFRS adoption. By examining the fees incurred by firms for the statutory audit of their financial statements following the adoption of IFRS, we quantify a significant and directly observable cost incurred by all firms. The notable variation of the net benefits of IFRS adoption observed in prior literature articulates the need to separate, and more fully understand, the costs associated with harmonization. Moreover, the recent concerns voiced by U.S. executives over the costs expected to arise from the potential U.S. adoption makes this study a timely exercise.

Using a comprehensive dataset of all publicly traded Australian companies, we estimate an economy-wide increase in the mean level of audit costs of 23 percent in the year of IFRS transition, varying with firm size and firm IFRS exposure. We estimate an abnormal IFRS-related increase in audit costs of 8 percent (i.e., beyond normal yearly fee increases). Further analysis also suggests a *fixed* component in the costs associated with IFRS adoption being borne by the smallest firms. Consequently, we find that small firms exhibit disproportionately larger increases in audit fees around the adoption of IFRS relative to large firms.

Our survey of professional auditors at a Big 4 accounting firm shows that auditors believe that certain aspects of the new IFRS reporting requirements (i.e., share-based incentive payments, financial instruments including hedge accounting, and impairment of goodwill and other intangible balances) require greater auditor effort and expertise to ensure adequate compliance. Constructing a firm-specific score of IFRS exposure based on our survey results, we confirm that the firms with the greatest exposure to these standards incur greater increases in audit fees in the year of adoption.

This study highlights one aspect of the costs associated with a mandatory regime change: the adoption of IFRS. Although we acknowledge that there are other internal implementation costs that are likely to be larger in magnitude than audit fees, an analysis of these costs is beyond the scope of this study. The discussion and results presented in this study have several implications for regulators and companies that are currently undertaking convergence projects throughout the world. Given the significant cost burden to small companies suggested by our results, local regulators and

the IASB may want to consider how to reduce the stress on small firms, as was done with the newly issued IFRS requirements for SMEs. Given the impending decision by the SEC about whether to permit a full-scale U.S. adoption of IFRS, current evidence on the costliness of IFRS and the specific provisions that are most costly and disruptive will provide valuable insight for future adopters and auditors to better tailor their transition programs.

REFERENCES

Andrews, B. 2005. Standards deviation. *Business Review Weekly* (February): 86.

Armstrong, C. S., M. E. Barth, A. D. Jagolinzer, and E. J. Riedl. 2010. Market reaction to the adoption of IFRS in Europe. *The Accounting Review* 85: 31–61.

Australian Accounting Standards Board (AASB). 2004a. *Share-Based Payment*. Australian Accounting Standard AASB 2. Melbourne, Australia: Australian Accounting Standards Board.

Australian Accounting Standards Board (AASB). 2004b. *Income Taxes*. Australian Accounting Standard AASB 112. Melbourne, Australia: Australian Accounting Standards Board.

Australian Accounting Standards Board (AASB). 2004c. *Employee Benefits*. Australian Accounting Standard AASB 119. Melbourne, Australia: Australian Accounting Standards Board.

Australian Accounting Standards Board (AASB). 2004d. *Impairment of Assets*. Australian Accounting Standard AASB 136. Melbourne, Australia: Australian Accounting Standards Board.

Australian Accounting Standards Board (AASB). 2004e. *Financial Instruments: Presentation and Disclosure*. Australian Accounting Standard AASB 132. Melbourne, Australia: Australian Accounting Standards Board.

Australian Accounting Standards Board (AASB). 2004f. *Financial Instruments: Recognition and Measurement*. Australian Accounting Standard AASB 139. Melbourne, Australia: Australian Accounting Standards Board.

Australian Accounting Standards Board (AASB). 2004g. *Intangibles*. Australian Accounting Standard AASB 138. Melbourne, Australia: Australian Accounting Standards Board.

Australian Institute of Company Directors (AICD). 2005. Submission to the Parliamentary Joint Committee on Corporations and Financial Services. Available at: http://www.companydirectors.com.au/Director-Resource-Centre/Policy-on-director-issues/Policy-Submissions/2005/Australian-Implementation-of-International-Financial-Reporting-Standards-IFRS

Ball, R. 2006. IFRS: Pros and cons for investors. *Accounting and Business Research* 36: 5–27.

Callao, S., J. I. Jarne, and J. A. Lainez. 2007. Adoption of IFRS in Spain: Effect on the comparability and relevance of financial reporting. *Journal of International Accounting, Auditing and Taxation* 16 (2): 148–178.

Charles, S. L., S. M. Glover, and N. Y. Sharp. 2010. The association between financial reporting risk and audit fees before and after the historic events surrounding SOX. *Auditing: A Journal of Practice & Theory* 29 (1): 15–39.

Chong, F. 2006. Firms fume over IFRS info "overkill." *The Australian*. Available at: http://www.theaustralian.com.au/business/firms-fume-over-ifrs-info-overkill/story-e6frg8zx-1111112588685

Clarkson, P., C. Ferguson, and J. Hall. 2003. Auditor conservatism and voluntary disclosure: Evidence from the year 2000 systems issue. *Accounting and Finance* 41: 21–40.

Craswell, A., J. R. Francis, and S. Taylor. 1995. Auditor brand names and industry specializations. *Journal of Accounting and Economics* (November): 297–322.

Daske, H., L. Hail, C. Leuz, and R. Verdi. 2008. Mandatory IFRS reporting around the world: Early evidence on the economic consequences. *Journal of Accounting Research* 46: 1085–1142.

Diehl, K. A. 2010. The real cost of IFRS: The relationship between IFRS Implementation and audit, tax, and other auditor fees. *International Research Journal of Finance and Economics* 37: 96–101.

Ernst & Young. 2005. The impacts of IFRS on Australian companies: A study of the financial statement disclosures by Australia's top 100 listed companies. Available at: http://www.ey.com.au

Ferguson, A., J. R. Francis, and D. Stokes. 2003. The effects of firm-wide and office-level industry expertise on audit pricing. *The Accounting* Review 2: 429–448.

Ghosh, A., and R. Pawlewicz. 2009. The impact of regulation on auditor fees: Evidence from the Sarbanes-Oxley Act. *Auditing: A Journal of Practice & Theory* 28 (2): 171–197.

Goodwin, J., and K. Ahmed. 2006. The impact of international financial reporting standards: Does size matter? *Managerial Auditing Journal* 21: 460–475.

Goodwin, J., K. Ahmed, and R.A. Heaney. 2008. The effects of International Financial Reporting Standards on the accounts and accounting quality of Australian firms: A retrospective study. *Journal of Contemporary Accounting and Economics* 4 (2): 89–119.

Griffin, P. A., D. H. Lont, and Y. Sun. 2009. Governance regulatory changes, International Financial Reporting Standards adoption, and New Zealand audit and nonaudit fees: Empirical evidence. *Accounting and Finance Journal* 49: 697–724.

Hail, L., C. Leuz, and P. Wysocki. 2010. Global accounting convergence and the potential adoption of IFRS by the United States (Part I): Conceptual underpinnings and economic analysis. *Accounting Horizons* 24 (3): 355–394.

Horton, J., and G. Serafeim. 2010. Market reaction to and valuation of IFRS reconciliation adjustments: First evidence from the U.K. *Review of Accounting Studies* 15: 725–751.

Institute of Chartered Accountants in Australia (ICAA). 2005. *Submission to the Parliamentary Joint Committee on Corporations and Financial Services*. Available at: http://www.aph.gov.au/Senate/committee/

Institute of Chartered Accountants for England & Wales (ICAEW). 2007. *EU Implementation of IFRS and the Fair Value Directive: Report for the European Commission*. London, U.K.: IACEW.

Jamal, K., R. H. Colson, R. J. Bloomfield, T. E. Christensen, S. R. Moehrle, J. Ohlson, S. H. Penman, G. J. Previts, T. L. Stober, S. Sunder, and R. L. Watts. 2009. *A Research Based Perspective on SEC's Proposed Rule on Roadmap for Potential Use of Financial Statements Prepared in Accordance With International Financial Reporting Standards (IFRS) by U.S. Issuers*. Working paper, Johnson School Research Paper Series No. #21-09. Available at: http://ssrn.com/abstract=1381292

Jermakowicz, E. K., and S. Gornik-Tomaszewski. 2006. Implementing IFRS from the perspective of EU publicly traded companies. *Journal of International Accounting, Auditing and Taxation* 15 (2): 170–196.

Jermakowicz, E. K., J. Prather-Kinsey, and I. Wulf. 2007. The value relevance of accounting income reported by DAX-30 German companies. *Journal of International Financial Management and Accounting* 18 (3): 151–191.

Johnson, S. 2009. Guessing the costs of IFRS conversion. *CFO Magazine*. Available at: http://www.cfo.com/article.cfm/13399306

Jubb, C. 2005. *Transition to IFRS: Listed Companies Expected Policy Impacts as Revealed by AASB 1047 Disclosures*. Working paper, Institute of Chartered Accountants in Australia, Deakin Business School.

Judge, G. G., R. C. Hill, W. E. Griffiths, H. Lutkepohl, and T. Lee. 1987. *Theory and Practice of Econometrics*. Second Edition. New York, NY: Wiley.

Kim J., X. Liu, and L. Zheng. 2010. *Does Mandatory IFRS Adoption Impact Audit Fees? Theory and Evidence*. Working paper, The University of Hong Kong.

Kittney, D., and F. Buffini. 2006. Mergers robbed of goodwill under IFRS. *Financial Review* (August): 11.

Lin, H. L., and A. R. Yen. 2009. *The Effects of IFRS Adoption on Audit Fees for Listed Companies in China*. Working paper, National Chiao-Ton University, Northeastern Illinois University.

Love, V. J., and J. H. Eickemeyer. 2009. IFRS and accountants' liability. *CPA Journal* 79 (4): 54–59.

Lynch, D., and A. Fabro. 2005. SMEs set to win relief on IFRS reporting deadline. *Financial Review* (February): 47.

Lyon, J. D., and M. W. Maher. 2005. The importance of business risk in setting audit fees: Evidence from cases of client misconduct. *Journal of Accounting Research* 43: 133–151.

PricewaterhouseCoopers LLP (PwC). 2008. *2007 Securities Litigation Study*. New York: NY: PwC Advisory Crisis Management.

Schadewitz, H., and M. J. Vieru. 2008. Impact of IFRS transition complexity on audit and nonaudit fees: Evidence from small and medium-sized listed companies in Finland. Available at: http://ssrn.com/abstract=967314

Securities and Exchange Commission (SEC). 2008. *Roadmap for the Potential Use of Financial Statements Prepared in Accordance with International Financial Reporting Standards by U.S. Issuers*. Available at: http://www.sec.gov/rules/proposed/2009/33-9005.pdf

Segarra, M. 2011. Small companies criticize switch to IFRS. *CFO Magazine*. Available at: http://www.cfo.com/article.cfm/14588474

Simunic, D. 1980. The pricing of audit services: Theory and evidence. *Journal of Accounting Research* (Spring): 161–190.

Van Garderen, K. J. 2001. Optimal prediction in loglinear models. *Journal of Econometrics* 104: 119–140.

Webb, R. 2006. Brace yourselves: IFRS will be bumpy. *Financial Review* (January): 26.

APPENDIX A

SUMMARY OF SIGNIFICANT CHANGES DUE TO IFRS

Accounting Standard Description	Significant Differences
Share-Based Payments (IFRS 2/AASB2)	Under IFRS 2 (AASB 2), companies are now required to recognize the expense relating to share-based payments on the face of the financial statements at fair value. Companies are required to estimate the fair value of employee stock options as of grant date and expense the value over the vesting period. However, in certain circumstances, the standard requires the use of subjective option-pricing models in determining fair values, which is problematic and complex. Recent academic literature has found that material adjustments stemming from the current and retrospective implementation of IFRS affect approximately 45 percent of Australian companies (Goodwin et al. 2008).
Income Taxes (IAS 12/AASB 112)	Previous AGAAP income tax expense was calculated by reference to the accounting profit after allowing for permanent differences, with deferred tax not recognized in relation to amounts recognized directly in equity. IFRS (Australian equivalent: AASB 112) now mandates the use of the "balance sheet" method for tax effect accounting where temporary differences are identified for each asset and liability. Accordingly, deferred tax balances are recognized on the balance sheet when differences arise between the carrying value of assets and liabilities, and their tax base. Revaluations due to fair value adjustments require the recognition of the attached tax effects. Thus, the inherent risks associated with fair value measurements directly affect the amount of deferred taxes recognized on the face of the financial statements. Prior literature suggests that more than 50 percent of the top 100 listed companies were significantly affected by this change (Ernst & Young 2005). Goodwin et al. (2008) find that 33 percent of ASX-listed companies reported material adjustments relating to these requirements.
Employee Benefits (IAS 19/AASB 119)	IFRS requires the recognition of any net surplus (deficit) of the plan funds as an asset (liability). Companies are required to obtain independent actuarial valuations of the fair value of plan assets and liabilities and then determine the extent of a net surplus or deficit, with significant disclosure requirements over plan details and future projections. While prior literature has demonstrated significant financial statement impacts of the application of IFRS over employee benefits (AASB 119) on a number of Australian companies (Ernst & Young 2005; Goodwin et al. 2008), companies and auditors are able to rely on independent actuarial valuations. Therefore, much of the subjectivity surrounding these calculations has been removed.
Impairment (IAS 36/AASB 136)	IFRS requires assets subject to impairment, primarily indefinite-life intangibles such as goodwill, to be held at the recoverable amount on the face of the financial statements, where the recoverable amount is defined as the higher of fair value less costs involved with sale or value in use determined by discounted cash flow methods. This standard introduces the mandatory requirement to discount future cash flows in the measurement of recoverable amounts. This effectively reduces the policy choices available under previous GAAP, and introduces significant subjectivity over inputs in impairment testing.

(continued on next page)

APPENDIX A (continued)

Accounting Standard Description	Significant Differences
Business Combinations (AASB 3)	IFRS requires identification of specific intangible assets and allocation of goodwill to cash-generating units expected to benefit from synergy, which results in an adjustment to purchase price allocations. Previous practice was to take the entire amount to goodwill. Furthermore, guidance under IAS 136 required firms to change their accounting policy to cease amortizing goodwill and undertake impairment testing annually, or when indicators of impairment are present. As such, prior amounts of amortization are reversed before opening IFRS balances are tested for impairment.
Intangibles (IAS 38/AASB 138)	Recognition under IFRS does not extend to internally generated intangibles. Thus, companies are forced to de-recognize any previously recognized internally generated intangibles. Moreover, recognized intangibles are now subject to impairment testing. The new IFRS requirements do not permit research expenditures to be capitalized, and the capitalization of development costs is now subject to principles-based recognition criteria; e.g. the ability to demonstrate the technical feasibility of developing an asset available for use or sale and the probability of generating future economic benefits. However, the precise guidance as to what constitutes technical feasibility remains vague, allowing increased management discretion in the capitalization of these costs. Purchased and internally generated software classified as intangible assets requires a reclassification from PPE. While these amounts are significant, they are of little consequence; hence, our data-collection procedures have not included these.
Financial Instruments (IAS 32 and 39/ AASB 132 and 139)	The requirements of IFRS provisions over financial instruments (AASB 132/139) are substantial relative to previous local GAAP. Measurement is now predominantly fair value-based, requiring complex and subjective estimates in the absence of active and liquid financial markets, and extensive disclosures concerning measurement assumptions and financial risk profiles. For example, AASB 139 (AASB 2004f) requires derivatives not meeting hedging criteria to be measured at fair value through the income statement and recognized on the balance sheet. Additionally, IFRS requirements over hedged instruments are also significant, with companies now required to perform extensive tests over hedge effectiveness and to ensure adequate documentation of relevant hedged relationships. Derivatives that don't comply with hedge accounting, hence deemed ineffective, must record all changes in fair value of the instruments directly through income. An instrument designated as an effective cash flow hedge has changes in fair value recognized directly to equity, with the ineffective portion flowing directly through income. When the underlying transaction being hedged affects the income statement, the deferred fair value changes are transferred from the hedge reserve into equity. Prior literature released by the accounting profession has reported that AASB 132/139 is proving the most difficult and onerous for companies in terms of their financial reporting and documentation obligations (Ernst & Young 2005).

This appendix provides a brief summary and description of the significant changes embodied in the new IFRS relative to existing Australian GAAP. Based on a review of the extant literature, we identify six key standards as having the most impact on firms' transitional process: *Share-Based Payments* (IFRS 2/AASB 2), *Income Taxes* (IAS12/AASB 112), *Employee Benefits* (IAS 19/AASB 119), *Impairment* (IAS 36/AASB 136), *Intangibles* (IAS 38/AASB 138), and *Financial Instruments* (IAS 32/39; AASB 132/139). Note that the standard numbers in parentheses indicate the standard as issued by the IASB and the corresponding Australian equivalent as issued by the AASB.

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 463–498

American Accounting Association
DOI: 10.2308/accr-50318

Managerial Ability and Earnings Quality

Peter R. Demerjian
Emory University

Baruch Lev
New York University

Melissa F. Lewis
University of Utah

Sarah E. McVay
University of Washington

ABSTRACT: We examine the relation between managerial ability and earnings quality. We find that earnings quality is positively associated with managerial ability. Specifically, more able managers are associated with fewer subsequent restatements, higher earnings and accruals persistence, lower errors in the bad debt provision, and higher quality accrual estimations. The results are consistent with the premise that managers can and do impact the quality of the judgments and estimates used to form earnings.

Keywords: *managerial ability; managerial efficiency; earnings quality; accruals quality.*

Data Availability: *Data are publicly available from the sources identified in the text.*

I. INTRODUCTION

We examine the relation between managerial ability and earnings quality. We anticipate that superior managers are more knowledgeable of their business, leading to better judgments and estimates and, thus, higher quality earnings.[1] Alternatively, the benefit

We thank two anonymous reviewers, Asher Curtis, Patty Dechow, Ilia Dichev, Weili Ge, Marlene Plumlee, Phil Shane, Terry Shevlin, Wayne Thomas (editor), Ben Whipple, and workshop participants at the 2010 Kapnick Accounting Conference at the University of Michigan, the 2010 Western Region Meeting, the 2010 FEA Conference, Emory University, Florida International University, University of Illinois, The University of Texas at Arlington, and the Wednesday group at The University of Utah for their helpful comments. An earlier version of this paper benefited from comments by Carol Anilowski Cain, Brian Cadman, Asher Curtis, Patty Dechow, Paul Michaud, Venky Nagar, Larry Seiford, Ram Venkataraman, Norman White, and workshop participants at the 2006 AAA Annual Meeting, Harvard University, University of California, Berkeley, and Indiana University.

Editor's note: Accepted by Wayne Thomas.

Submitted: July 2010
Accepted: August 2012
Published Online: October 2012

[1] Following Dechow and Schrand (2004), we consider high-quality earnings to be those that accurately reflect companies' current operating performance, and assume that managers take a shareholder/analyst perspective when determining the desired attributes of earnings (see Dechow et al. 2010). As we discuss further below, the managers we focus on are CFOs and their delegates, as we are interested in the accuracy of accrual estimations.

of higher quality earnings may not be sufficient to warrant the time and attention of skilled management, especially if the variance of feasible estimates is small, in which case we may not find an association between managerial ability and earnings quality.[2]

While the archival literature in the area of earnings quality has largely focused on firm-specific characteristics, such as firm size and board independence (Dechow and Dichev 2002; Klein 2002), we examine the manager-specific aspect of earnings quality. Our study is in the vein of Bertrand and Schoar (2003), who find that managers have an effect on firm choices such as acquisitions or research and development expenditures; Aier et al. (2005), who document that CFOs with more accounting expertise have fewer restatements; and Francis et al. (2008), who document that earnings quality varies inversely with CEO reputation.[3]

Our main measure of managerial ability (hereafter, the MA-Score) is developed in Demerjian et al. (2012), although we perform robustness checks using historical returns, media citations, and manager fixed effects (e.g., Fee and Hadlock 2003; Milbourn 2003; Francis et al. 2008; Dyreng et al. 2010). Demerjian et al. (2012) first estimate total firm efficiency, where efficient firms are those that generate more revenues from a given set of inputs. Total firm efficiency is influenced by both the manager, who can, to varying degrees, predict future demand and understand industry trends, and the firm, because (for example) managers in larger firms can negotiate better terms. Thus, Demerjian et al. (2012) partition total efficiency between the firm and the manager, and verify that the component attributed to the manager is associated with a variety of characteristics, including managerial pay and the price reaction to management departures from the firm.[4] Prior research is limited to measures such as media coverage and historical returns, which are difficult to attribute solely to the manager versus the firm (Francis et al. 2008), or manager fixed effects, where there is evidence of a manager-specific effect, but the quantifiable effect is limited to managers who switch firms (e.g., Bertrand and Schoar 2003; Bamber et al. 2010; Ge et al. 2011). The MA-Score allows us to better distinguish the effect of the manager from the effect of the firm and to retain an ordinal ranking of quality for a large sample of firms.[5]

We expect a more able management team to estimate accruals more accurately. For example, we expect more able managers to be more knowledgeable of their client base and macro-economic conditions when estimating bad debt expense, be more knowledgeable of the expected future benefits of recorded assets, and to be more able to understand and apply complex standards (e.g., McNichols 2002; Plumlee and Yohn 2010).

---

[2] Costs of poor earnings quality include higher cost of capital (Francis et al. 2004) and economically significant negative price reactions to the announcement of earnings restatements (Palmrose et al. 2004).

[3] We hypothesize a positive relation between ability and earnings quality, which is opposite to the relation documented in Francis et al. (2008). Francis et al. (2008) measure CEO reputation with the number of articles mentioning the executive and document a negative association between the number of news articles pertaining to the company's CEO and earnings quality based on the Dechow and Dichev (2002) accruals quality measure. We find that this negative relation in Francis et al. (2008) appears to be due, at least in part, to measurement error in the accruals quality measure (see Section IV).

[4] As we clarify in the following sections, Demerjian et al. (2012) estimate total firm efficiency using data envelopment analysis, a type of frontier analysis that measures relative efficiency (see also Knechel et al. 2009). They then remove identifiable firm characteristics, such as size, that affect the firm's relative efficiency but are unlikely to be a direct result of the quality of management. They attribute the unexplained portion of total firm efficiency to the management team. They document that their measure outperforms existing measures of ability such as historical stock returns and media citations.

[5] The ability score is for the management team. In our setting we would like to determine the quality of CFOs and their delegates, as we focus on the estimation of accruals, whereas CEOs are more focused on the overall strategy of the firm. Although we cannot disentangle the ability score by CEO and CFO, the ability score does encompass CFOs and their delegates, whereas media citations are, by definition, focused on the CEO. We also identify, in Section V, CFOs who switch firms within our sample to document CFO-specific fixed effects on accruals quality and correlate the CFOs' scores from their old firms with the accruals quality after their arrival in their new firms.

We consider four measures of earnings quality: the existence of an earnings restatement (Anderson and Yohn 2002), the persistence of earnings (Richardson et al. 2005), errors in the bad debt provision (McNichols and Wilson 1988), and the extent to which accruals map into cash flows (Dechow and Dichev 2002).[6] In general, we find that earnings quality is positively associated with managerial ability. This finding is consistent with the premise that more capable managers are better able to estimate accruals, which results in a more precise measure of earnings.

We contribute to both the earnings quality literature and the managerial ability literature by establishing a positive and significant relation between managerial ability and earnings quality, which suggests a means of improving earnings quality. Many of the factors associated with earnings quality, such as firm size, industry, or operating cycle, result from strategic goals and competitive advantages of the firm. Consequently, it may not be advantageous to improve earnings quality by changing these characteristics. In contrast, we conclude that given the set of earnings-estimation challenges resulting from the firm's operating decisions, higher ability managers will be able to better deal with these complexities and report higher quality earnings relative to similar firms operating in similar environments. This finding is important for board members when considering the costs and benefits of managers because managerial ability affects not only the operations of the firm, but also the quality of its reported earnings and, in turn, its share-price attributes and litigation exposure.

Our results also help reconcile the counter-intuitive prior findings that more reputable managers are associated with lower accruals quality (Francis et al. 2008). We find that this relation appears to be due, in part, to the impact of fundamental firm characteristics on the accruals quality measure. As noted by Dechow and Dichev (2002), the accruals quality measure is influenced by both intentional and unintentional errors, and many unavoidable unintentional errors stem from complex operating environments (Dechow and Dichev 2002). Following the spirit of Ball and Shivakumar (2006), we modify the Dechow and Dichev model to allow the coefficients of the model to vary with firm fundamentals. This allows the model to incorporate variations in the expected relation between accruals and cash flows across firms (see also Wysocki 2009). Our modifications allow us to document a positive and significant association between managerial ability and accruals quality.

In the next section, we develop our hypothesis with a review of the literature. In Section III we describe our sample, test variables, and descriptive statistics. In Section IV we present the main results, and in Section V we consider alternative ability measures and conduct a change analysis for a subset of managers in our sample who switch firms. We conclude the study in the final section.

II. HYPOTHESIS DEVELOPMENT AND RELATED LITERATURE

Earnings quality is an important characteristic of financial reports that affects the efficient allocation of resources. Because earnings are the main input to investors' and analysts' valuation models, firms with poor earnings quality tend to have higher costs of capital (e.g., Francis et al. 2004) and those experiencing restatements or SEC enforcement actions tend to experience an economically significant negative price reaction to the announcement (Feroz et al. 1991; Palmrose et al. 2004). Following Dechow and Schrand (2004), we define high-quality earnings to be those that accurately reflect companies' current operating performance.

We expect managers' ability to form accurate judgments and estimates to vary across individuals. We expect more able managers to be more knowledgeable about the firm and the industry, as well as to be better able to synthesize information into reliable forward-looking

[6] We use the term "earnings quality" to capture the general construct of higher quality reported earnings, while we use the term "accruals quality" to discuss the Dechow and Dichev measure of earnings quality based on the mapping of accruals to cash flows.

estimates with which to report higher quality earnings (e.g., Libby and Luft 1993). Specifically, we expect accruals estimated by high-ability managers to be more accurate than those estimated by low-ability managers. For example, consider the allowance for bad debt estimate. A less able manager might apply the historical rate of bad debt for the firm, while a more able manager might adjust the historical rate by considering the macro-economic and industry trends, as well as changes in the firm's customer base. Similarly, we expect more able managers to report more accurate and justifiable depreciation rates, fair values, and other accrual estimates. Thus, holding the firm constant, we expect a more able manager to report higher quality earnings.

H1: Managerial ability is positively associated with earnings quality.

It is possible that the majority of the variation in earnings quality is driven by innate firm characteristics that managers cannot affect, in which case we will not find an association between the ability of managers and the quality of earnings. It is also possible that the benefits to the incremental improvement in earnings quality resulting from the intervention by an able manager do not exceed the cost of that manager's time, in which case, again, we will not find an association between the manager's ability and the firm's earnings quality.

To date, the bulk of the literature on earnings quality has examined firm-specific characteristics. For example, Dechow and Dichev (2002) document that earnings quality is poorer for firms that are smaller, are experiencing losses, have greater sales and cash flow volatility, and have longer operating cycles. Each of these innate firm characteristics makes accruals more difficult to estimate. In addition to these innate characteristics, earnings quality has been found to vary with firm infrastructure, such as internal control quality (Doyle et al. 2007; Ashbaugh-Skaife et al. 2008) and monitors such as auditors (e.g., Becker et al. 1998) and boards (Klein 2002).

With respect to the effect of managers on the firm, Bamber et al. (2010) find that individual managers appear to have particular "styles" that are associated with their propensity to issue guidance and the characteristics of the resulting guidance (e.g., the precision of the guidance). In a similar vein, both Ge et al. (2011) and DeJong and Ling (2010) examine manager fixed effects on certain financial reporting policies and, similar to Bamber et al. (2010), document that individual managers matter because firms' accounting and disclosure policies vary with manager fixed effects. As previously noted, this approach allows researchers to document a manager-specific effect, but it is constrained to managers who switch employers among the sample firms. Fixed effects are also usually limited to larger firms as executives within smaller firms often switch to private firms (Ge et al. 2011).

Of particular relevance for our study, both Aier et al. (2005) and Francis et al. (2008) examine whether earnings quality varies with managerial characteristics. Aier et al. (2005) document an association between CFO expertise (years worked as CFO, experience at another company, advanced degrees, and professional certifications) and restatements, concluding that firms employing CFOs who have greater expertise have fewer restatements. Francis et al. (2008) examine the relation between earnings quality and CEO reputation, measured by the number of business press articles mentioning each CEO. They find a *negative* relation between CEO reputation and earnings quality. They conclude that "boards of directors hire specific managers due to the reputation and expertise these individuals bring to managing the more complex and volatile operating environments of these firms" Francis et al. (2008).

In sum, there is mixed archival evidence on the impact of managers on earnings quality. Although there is some evidence that managers with greater expertise are associated with fewer earnings restatements, Francis et al. (2008) document that more reputable managers are associated with lower earnings quality. The latter association is consistent with some firms having less predictable, and thus lower quality, earnings by the nature of their business (Dechow and Schrand 2004; LaFond 2008) and these firms hiring better managers. Although it is more difficult to estimate

accruals within certain environments, such as more volatile firms, we expect that better managers can estimate accruals more accurately for a given environment, for example, within loss firms.

We consider four earnings quality measures. The first is earnings restatements, which are *de facto* evidence of inaccurate earnings (Dechow et al. 2010).[7] The second is earnings persistence, where we partition earnings into accrual and cash flow components to examine accruals persistence more directly (Sloan 1996; Richardson et al. 2005). Our third earnings quality measure is the accuracy of the bad debt provision (McNichols and Wilson 1988). Finally, we examine the mapping of working capital accruals into cash from operations, based on Dechow and Dichev (2002).

Each of these measures is affected by both unintentional errors and intentional errors, and more able managers may be more likely to introduce intentional errors, either to signal their private information about the firm or to extract perquisites from the firm and the shareholders. We focus on earnings quality measures that capture estimation errors in accruals, but do not attempt to distinguish between intentional and unintentional errors.[8]

III. DATA, VARIABLE DEFINITIONS, AND DESCRIPTIVE STATISTICS

We obtain our data from the 2010 Annual Compustat File for the bulk of our earnings quality variables and controls; from CRSP to form historical returns, an alternate managerial ability measure; from ExecuComp to track CFOs across firms; from IRRC to obtain board independence data; and from Audit Analytics for recent years of restatements and internal control opinions. We also obtain several datasets made available by researchers, including managerial ability from Demerjian et al. (2012), media citations from Baik et al. (2011), restatements from Hennes et al. (2008) and Plumlee and Yohn (2010), and internal control quality data from Doyle et al. (2007).

We begin with all firms with managerial ability data and at least one of our earnings quality variables. We then exclude firm-years with acquisition activity in excess of 5 percent of assets, resulting in a maximum of 78,423 firm-year observations from 1989–2009. The period begins with 1989 because 1988 is the first year for which firms widely reported cash flow statements, and the Dechow and Dichev accruals quality variable requires one year of historical cash flow data. The sample ends in 2009 because our earnings quality variables described in the following section require at least one year of future realizations. We exclude firms with material acquisition activity as it could unduly affect both the measure of managerial ability and our earnings quality measures.

---

[7] Some erroneous estimates of accruals are corrected prospectively, in that adjustments are made going forward, but prior earnings are not restated. In these cases, restatements offer an arguably weaker measure of earnings quality than the other measures we examine. Many erroneous judgments and estimates, however, are retrospectively restated. For example, when marking-to-market, managers must determine whether to rely on traded prices or other valuations when traded prices do not appear to be reflective of the fair value. Further, managers must determine whether to write-down assets based on their judgment of whether the assets will be realized as future benefits. In their appendix, Plumlee and Yohn (2010) highlight restatements for both of these settings in which managers' judgments were *ex post* deemed to be wrong. We corroborate in subsequent analyses that a large number of restatements are attributable to managerial judgments and estimates, and that these are the types of restatements that are associated with managerial ability.

[8] Because we examine a broad cross-section, we do not expect to find strong evidence of earnings management, on average (Dechow and Skinner 2000). It is possible that better managers are more likely to smooth earnings or otherwise use earnings management as a signaling mechanism (Tan and Jamal 2006), but it is also possible that they can more effectively use earnings management to extract personal benefits. Our measure of managerial ability is based on efficiency, and does not incorporate ethical considerations (Kim et al. 2012). In the event of earnings management, however, we expect more able managers to be better able to manage earnings *successfully*, for example, accelerating sales only if they know there will be sufficient sales in the next period to cover the acceleration, thereby avoiding large accrual reversals and restatements. We leave a direct examination of the interaction between managerial ability and earnings management for future research.

American Accounting Association

Variable Definitions

Managerial Ability Measure

Our main measure of managerial ability, the MA-Score, is developed by Demerjian et al. (2012), who generate an estimate of how efficiently managers use their firms' resources. All firms use capital, labor, and innovative assets to generate revenues. High-quality managers will generate a higher rate of output from given inputs than lower quality managers, for example by applying superior business systems and processes, such as supply chains and compensation systems.

Demerjian et al. (2012) use data envelopment analysis (DEA) to estimate firm efficiency within industries, comparing the sales generated by each firm, conditional on the following inputs used by the firm: Cost of Goods Sold, Selling and Administrative Expenses, Net PP&E, Net Operating Leases, Net Research and Development, Purchased Goodwill, and Other Intangible Assets.[9] Thus, the measured resources reflect tangible and intangible assets, innovative capital (R&D), and other inputs that are not reported separately in the financial statements, such as labor and consulting services, but whose costs are included in cost of sales and SG&A. We provide the motivation and definition for each of these variables in Appendix A.

Demerjian et al. (2012) use DEA to solve the following optimization problem:

$$\max_{\boldsymbol{v}}\boldsymbol{\theta} = \frac{Sales}{v_1CoGS + v_2SG\&A + v_3PPE + v_4OpsLease + v_5R\&D + v_6Goodwill + v_7OtherIntan}.$$

The optimization finds the firm-specific vector of optimal weights on the seven inputs, $\boldsymbol{v}$, by comparing each of the input choices of an individual firm to those of the other firms in its estimation group. The efficiency measure that DEA produces, $\boldsymbol{\theta}$, takes a value between 0 and 1, reflecting constraints in the optimization program. Observations with a value of 1 are the most efficient and the set of firms with efficiency equal to 1 trace a frontier through the efficient set of possible input combinations. Observations with efficiency measures less than 1 fall below the frontier. A firm's DEA score indicates the degree to which the firm is efficient. A firm with a score of less than 1 would need to reduce costs or increase revenues to achieve efficiency.

The efficiency measure generated by the DEA estimation is attributable to both the firm and the manager, similar to other measures of managerial ability such as historical returns and media coverage. For example, a more able manager will be better able to predict trends, regardless of the size of the firm, while a manager in a larger firm will, on average, be better able to negotiate terms with suppliers, regardless of his or her quality. Demerjian et al. (2012) therefore modify the DEA-generated firm efficiency measure by purging it of key firm-specific characteristics expected to aid or hinder management's efforts, including firm size, market share, positive free cash flow, and firm age, which aid management, and complex multi-segment and international operations, which challenge management.[10] They estimate the following Tobit regression model by industry:

---

[9] DEA is a frontier analysis that calculates efficiency as the ratio of weighted outputs to weighted inputs. DEA uses an optimization program to determine the firm-specific optimal or "implicit" weights on the inputs and outputs. The implicit weights capture the efficiency of the firm based on the selected inputs and outputs, allowing the optimal mix of inputs and outputs to vary by firm. This differs from other efficiency measures, such as ROA or ROE, that require an explicit set of weights, generally equal to 1. Since Demerjian et al. (2012) have only one output, sales, its weight is standardized to 1 across observations. For the general DEA model, please see Appendix A.

[10] To the extent that managers also affect some of the independent variables in Equation (1), such as free cash flow, the MA-Score is a conservative (understated) measure of managerial efficiency. We also supplement Demerjian et al.'s (2012) estimation by including risk as an additional independent variable, measured using leverage and beta. We control for risk because superior operating performance that results from riskier operations is not necessarily indicative of higher ability managers. We find similar results when we control for risk, indicating that the ability measure is not simply picking up firms with riskier operations.

$$\begin{aligned} \textit{Firm Efficiency} = {} & \alpha_0 + \alpha_1 Ln(\textit{Total Assets}) + \alpha_2 \textit{Market Share} + \alpha_3 \textit{Positive Free Cash Flow} \\ & + \alpha_4 Ln(\textit{Age}) + \alpha_5 \textit{BusinessSegmentConcentration} \\ & + \alpha_6 \textit{Foreign Currency Indicator} + \textit{Year Indicators} + \varepsilon. \end{aligned} \quad (1)$$

The residual from the estimation is the MA-Score, which we attribute to the management team and rely on as our main measure of managerial ability (*MgrlAbility*).[11] We create decile ranks of *MgrlAbility* by year and industry to make the score more comparable across time and industries and to mitigate the influence of extreme observations. Untabulated analyses indicate that results are similar using a continuous variable.

Demerjian et al. (2012) corroborate this measure by performing a number of validity tests. First, the MA-Score is strongly associated with manager fixed effects, suggesting it reflects manager characteristics, not just firm characteristics omitted from Equation (1). Second, they document a negative (positive) stock price reaction when high-ability (low-ability) CEOs announce they are leaving the firm. Third, they find that replacing CEOs with new CEOs of higher ability (lower ability) is associated with improvements (declines) in subsequent firm performance. The MA-Score is also positively correlated with CEO pay and historical returns and outperforms historical returns, historical ROA, compensation, tenure, and media citations in each of their tests. Together, their validity tests provide strong evidence that the MA-Score reflects managerial talent that is distinct from the firm.

While the MA-Score is our main measure of managerial ability, it has several possible sources of error in its measurement. First, Demerjian et al. (2012) acknowledge that in the first-stage DEA estimation the inputs and output are measured with noise. For example, accounting variables, such as sales and cost of goods sold, can be manipulated by management and can be measured differently across firms, and some variables of interest, such as advertising or purchased research and development are not available for most firm-year observations. To the extent that certain input data are not available or are measured with error, the production function underlying the DEA estimation is potentially incomplete or inaccurate. In addition, the second-stage estimation attributes any firm efficiency outside of the set of the identified firm features, which are the explanatory variables in Equation (1), to managerial ability. If the set of firm features is incomplete, then the measure may overstate managerial ability by attributing efficiency inherent to the firm to the manager. Thus, in Section V we also consider alternative measures of managerial ability, including media citations, historical stock returns, and manager fixed effects.

Earnings Quality Measures

Dechow et al. (2010) note that there are a multitude of earnings quality measures used in the literature. To examine the impact of managers on accrual estimation, we select earnings restatements, earnings persistence, errors in the bad debt provision, and the mapping of accruals into cash flows as our four measures of earnings quality. We select these measures because increased correspondence between accruals and the associated economic activity likely reduces earnings restatements, increases earnings persistence, and lowers the likelihood of errors in accruals. Because we expect that better managers are able to report accruals that more closely correspond to the underlying economic activity, we expect the earnings quality metrics that are affected by judgments and accrual estimation to vary with managerial ability. For each of these measures, we consider earnings quality in year *t*+1 onward, i.e., in periods subsequent to year *t*, when managerial ability is measured. This reduces the likelihood that an economic shock

[11] The MA-Score data are available at https://community.bus.emory.edu/personal/PDEMERJ/Pages/Home.aspx.

concurrently affects both our measurement of ability and earnings quality. We discuss each earnings quality metric in greater detail in Section IV.[12]

Control Variables

Our main set of control variables is based on the firm-specific determinants of earnings quality noted in Dechow and Dichev (2002) and Hribar and Nichols (2007), including firm size, proportion of losses, sales volatility, cash flow volatility, and operating cycle. We also control for whether the company's auditor is a national audit firm, which is associated with earnings quality (Becker et al. 1998). Finally, we control for change in sales growth and abnormal returns to control for growth and economic shocks to performance, both of which could potentially impact our measures of managerial ability and earnings quality. We provide variable definitions and measurement periods in Table 1.

Descriptive Statistics

For each of our transformed variables (*MgrlAbility*, *Historical Ret*, *Media Citations*, *AQ* [and modifications thereof], *Firm Size*, and *Operating Cycle*), we present the untransformed variable for ease of interpretation in Table 1. By construction, managerial ability has a mean and median close to 0, as this is a residual from Equation (1). The five-year historical return has a mean of approximately 6 percent and, on average, CEOs are cited by the media approximately 44 times per year or 219 times over five years. Approximately 13 percent of firms experience a restatement in the next three years, and firm-specific earnings persistence averages approximately 0.23.[13] The error in the provision for bad debt as a percentage of sales (*BDE Error*) has a mean and median of 0.01. Mean (median) *AQ* is −0.03 (−0.03), similar to that in Francis et al. (2004) and Dechow and Dichev (2002), where we have multiplied the standard deviation by −1.

In Panel B of Table 1 we partition our earnings quality measures by managerial ability, where low-quality (high-quality) managers are those in the bottom (top) quintile of managerial ability, where quintiles are formed by industry-year. Historical returns are significantly higher among high-quality managers, consistent with Fee and Hadlock (2003) and Demerjian et al. (2012), although media citations are significantly lower for managers with higher ability, a relation we explore in Section V. Restatements are more prevalent among low-quality managers, median firm-specific earnings persistence is higher among high-quality managers, and errors in the provision for bad

---

[12] We do not consider the absolute value of discretionary accruals, earnings smoothness and benchmarking, as the relation between improved accruals estimation and these metrics is not clear. For example, abnormally high accruals may be high-quality accruals that are associated with cash flows, while abnormally low accruals may reflect extreme negative performance, which also reflects the underlying economics of the firm. Neither of these "abnormal" accruals provides information on the manager's ability to appropriately estimate accruals, as the measure does not incorporate *ex post* realizations. We do not consider timely loss recognition, as it is not apparent whether more or less timely loss recognition better reflects the underlying economics of the firm. As noted in Dechow et al. (2010), ERCs are a poor measure of earnings quality because much of the earnings information can be voluntarily disclosed prior to the earnings announcement. Finally, of the three external indicators of earnings quality—restatements, AAERs, and internal control disclosures—we consider only restatements. We do not consider AAERs because these tend to be more fraudulent than basic errors in estimation (Hennes et al. 2008). We do not consider internal control deficiencies as an outcome because the determinants of internal control problems are largely firm-specific, such as having adequate resources to establish and maintain these controls. The role of an able manager in the determination of strong internal controls is less clear, and does not speak to management's ability to estimate accruals.

[13] We identify the firm-specific persistence in order to have a stand-alone measure of persistence. The 0.23 firm-specific persistence coefficient is lower than the typical coefficient for cross-sectional persistence because it is a time-series, firm-specific coefficient, based on quarterly observations of earnings per share rather than the more traditional annual observations of ROA. When we estimate a cross-sectional regression by year, we find a mean earnings per share persistence coefficient of 1 and an ROA persistence coefficient of about 0.70.

TABLE 1

Descriptive Statistics

Panel A: Descriptive Statistics for the Full Sample (1989–2009)

Variable	n	Mean	Median	Std. Dev.	25%	75%
MgrlAbility[a]	78,423	0.00	−0.01	0.15	−0.09	0.07
Historical Ret[a]	40,871	0.06	−0.35	2.17	−1.00	0.50
Media Citations[a]	10,110	219.11	93.00	693.55	46.00	180.00
Restate	46,022	0.13	0.00	0.33	0.00	0.00
F.S. EarnPer	68,447	0.23	0.18	0.40	−0.02	0.48
BDE Error	1,124	0.01	0.01	0.05	0.00	0.01
AQ[a]	51,925	−0.03	−0.03	0.03	−0.05	−0.01
Modified $AQ_{LOSS\%}$[a]	52,316	−0.05	−0.03	0.05	−0.07	−0.02
Modified $AQ_{SI\_LOSS\%}$[a]	52,316	−0.05	−0.04	0.05	−0.07	−0.02
Modified $AQ_{NEGCFO\%}$[a]	52,317	−0.05	−0.04	0.05	−0.07	−0.02
TotalEarnQuality	17,128	0.91	1.00	0.56	0.55	1.33
ΔWC	78,423	0.01	0.01	0.12	−0.03	0.05
CFO	78,423	−0.01	0.06	0.32	−0.03	0.13
FirmSize[a]	78,423	1,134.66	84.80	4,645.46	17.92	423.29
Loss%	73,231	0.40	0.33	0.37	0.00	0.80
SalesVolatility	68,152	0.23	0.16	0.23	0.09	0.30
CashFlowVolatility	65,641	0.10	0.06	0.14	0.04	0.12
OperCycle[a]	77,444	158.72	112.65	281.68	68.01	174.53
$\Delta SalesGrowth$	67,849	−0.05	0.00	3.14	−0.15	0.31
AbnormalReturn	62,928	0.02	−0.11	0.73	−0.39	0.23
FutureEarnings	78,423	−0.12	0.02	0.73	−0.10	0.08

[a] For each of our transformed variables (*MgrlAbility*, *Historical Ret*, *Media Citations*, *AQ*, *Modified AQ*, *FirmSize*, and *OperCycle*), we present the untransformed variable for ease of interpretation.

Panel B: Accruals Quality Variables by Managerial Ability

	Lowest Quintile of *MgrlAbility*		Highest Quintile of *MgrlAbility*			
Variable	Mean	Median	Mean	Median	Diff. Mean	Diff. Med.
MgrlAbility[a]	−0.18	−0.17	0.20	0.18	***	***
Historical Ret[a]	−0.42	−0.65	0.79	0.09	***	***
Media Citations[a]	278.48	107.00	151.07	79.00	***	***
Restate	0.14	0.00	0.12	0.00	***	***
F.S. EarnPer	0.21	0.15	0.25	0.20		***
BDE Error	0.02	0.01	0.00	0.01	***	***
AQ[a]	−0.036	−0.025	−0.037	−0.028	***	***
TotalEarnQuality	0.82	0.89	0.93	1.00	***	***

*** Denotes a difference in the mean (median) under a t-test (Chi-square test) with a two-tailed p-value of less than 0.01.
[a] For each of our transformed variables (*MgrlAbility*, *Historical Ret*, *Media Citations*, *AQ*, *Modified AQ*, *Firm Size*, and *OperCycle*), we present the untransformed variable for ease of interpretation.
All continuous variables are winsorized at the extreme 1 percent. All variables are reported as of year *t* in this table only.

(continued on next page)

TABLE 1 (continued)

Panel C: Variable Definitions

Variable	Description	Definition
Ability Measures		
MgrlAbility	Managerial ability	The decile rank (by industry and year) of the MA-Score, which is managerial efficiency from Demerjian et al. (2012) in year *t*; the residual from Equation (1); see Appendix A.
Historical Ret	Historical return	The decile rank (by industry and year) of the five-year past value-weighted industry-adjusted return (year $t-4$, t) using monthly CRSP data.
Media Citations	Media citations	The decile rank (by industry and year) of the number of articles mentioning the CEO over the preceding five-year period (year $t-4$, t).
Earnings Quality Measures		
Restate	Restatement	An indicator variable that is equal to 1 if the firm announced a restatement in years $t+1$, $t+2$, or $t+3$, and 0 otherwise (available from 1997–2009).
Restate Judgments	Restatement relating to judgments and estimates	An indicator variable that is equal to 1 if the firm announced a restatement in years $t+1$, $t+2$, or $t+3$ classified as *standards-based* or *complexity-based* per Plumlee and Yohn (2010), and 0 otherwise (available from 2003–2006). Standards-based restatements are one of three types: (1) restatements stemming from lack of clarity in the standard; (2) restatements resulting from mistakes in judgment; and (3) restatements stemming from errors in applying complex rules. Complexity-based restatements are those resulting from the complexity of a transaction.
Restate Other	Restatement relating to items other than judgment	An indicator variable that is equal to 1 if the firm announced a restatement in years $t+1$, $t+2$, or $t+3$ not classified as a restatement relating to judgments and estimates as defined above, per Plumlee and Yohn (2010), and 0 otherwise (available from 2003–2006).
F.S. EarnPer	Firm-specific earnings persistence	The firm-specific time-series coefficient on earnings (per share) in a regression of one-quarter-forward earnings on current-quarter earnings. We estimate firm-specific quarterly earnings persistence over years $t+1$ through year $t+4$.

(continued on next page)

TABLE 1 (continued)

Variable	Description	Definition
BDE Error	Unexplained bad debt expense	The absolute value of the residual (φ_t) from Equation (4) where three industries are considered: printing and publishing, nondurable wholesale goods, and business services.
AQ	Standard deviation of accrual errors	The decile rank (by industry and year) of −1 × Standard Deviation (ε_{t+1}, ε_{t+2}, ε_{t+3}, ε_{t+4}), where ε_{t+n} is the residual from Equation (6) estimated by industry-year, where industries are defined per Fama and French (1997).
Modified $AQ_{Loss\%}$	Modified standard deviation of accrual errors	The decile rank (by industry and year) of −1 × Standard Deviation (ε_{t+1}, ε_{t+2}, ε_{t+3}, ε_{t+4}), where ε_{t+n} is the residual from Equation (8) estimated by industry and the quintile rank of *Loss%*, where ranks are assigned annually by industry. Industries are defined per Fama and French (1997).
Modified $AQ_{SI\_Loss\%}$	Modified standard deviation of accrual errors	The decile rank (by industry and year) of −1 × Standard Deviation (ε_{t+1}, ε_{t+2}, ε_{t+3}, ε_{t+4}), where ε_{t+n} is the residual from Equation (8) estimated by industry and the quintile rank of *Loss% Before Special Items*, where ranks are assigned annually by industry. Industries are defined per Fama and French (1997).
Loss% Before Special Items	Loss percentage before special items	The percentage of years reporting losses in net income (IBC) excluding the impact of special items over at least three of the last five years (*t*–4, *t*). We exclude the impact of special items by subtracting positive special items from IBC and adding back negative special items to IBC.
Modified $AQ_{NegCFO\%}$	Modified standard deviation of accrual errors	The decile rank (by industry and year) of −1 × Standard Deviation (ε_{t+1}, ε_{t+2}, ε_{t+3}, ε_{t+4}), where ε_{t+n} is the residual from Equation (8) estimated by industry and the quintile rank of *Negative CFO%* where ranks are assigned annually by industry. Industries are defined per Fama and French (1997).
Negative CFO%	Negative CFO percentage	The percentage of years reporting negative cash flows from operations over at least three of the last five years (*t*–4, *t*).
TotalEarnQuality	Earnings quality summation variable	The sum of three earnings quality variables: (1) the rank of estimation accruals quality (*Modified* $AQ_{Loss\%}$); (2) the rank of firm-specific earnings persistence; and (3) −1 × *Restate*. Thus, the variable ranges from a low of −1 to a high of 2.

(continued on next page)

TABLE 1 (continued)

Variable	Description	Definition
TotalEarnQuality2	Earnings quality summation variable	The sum of three earnings quality variables: (1) the rank of estimation accruals quality using the annual AQ metric (*Modified* $AQ_{Loss\%Ann}$); (2) the rank of firm-specific earnings persistence; and (3) $-1 \times$ *Restate*. Thus, the variable ranges from a low of -1 to a high of 2.
Control Variables		
FirmSize	Firm size	The natural log of the firm's assets (AT) reported at the end of year *t*.
SalesVolatility	Sales volatility	The standard deviation of sales [(SALE)/average assets (AT)] over at least three of the last five years (*t*–4, *t*).
CashFlowVolatility	Cash flow volatility	The standard deviation of cash from operations [(OANCF)/average assets (AT)] over at least three of the last five years (*t*–4, *t*).
OperCycle	Operating cycle	The natural log of the length of the firm's operating cycle, defined as sales turnover plus days in inventory [(SALE/360)/(average RECT) + (COGS/360)/(average INVT)] and is averaged over at least three of the last five years (*t*–4, *t*).
Loss%	Loss history	The percentage of years reporting losses in net income (IBC) over at least three of the last five years (*t*–4, *t*).
National Auditor	National auditor indicator	An indicator variable set equal to 1 for firms audited by national audit firms in year *t*, 0 otherwise.
ΔSalesGrowth	One-year change in % sales growth	The one-year change in sales growth defined as current year's sales growth ($\Delta SALE_t/SALE_{t-1}$) less prior year's sales growth ($\Delta SALE_{t-1}/SALE_{t-2}$).
AbnormalReturn	Abnormal return	One-year market-adjusted buy-and-hold return for year *t* where market returns are value-weighted.
PctInd	Board independence	The percentage of board members classified as independent based on IRRC's classification (available from 1996–2007).
ICW	Internal control weakness	An indicator variable for firms reporting material weaknesses in internal control (available from 2002–2007).
Other Variables		
FutureEarnings	Future net income	Future earnings (IBC) scaled by average total assets (AT).
Earnings	Net income	Earnings (IBC) scaled by average total assets (AT).

(continued on next page)

TABLE 1 (continued)

Variable	Description	Definition
ΔWC	Working capital accruals	The change in working capital scaled by average total assets, where working capital is defined as follows: [− (RECCH + INVCH + APALCH + TXACH + AOLOCH)].
ΔREV	Change in sales	Current year change in sales (SALE) scaled by average total assets (AT).
PPE	Property, plant, and equipment	Current year level of property, plant, and equipment (PPENT) scaled by average total assets (AT).
CFO	Cash from operations	Cash from operations (OANCF) scaled by average total assets (AT).
Accruals	Accruals	Accruals (scaled by average total assets (AT)), where Accruals = Earnings − CFO.

Subscripts in Panel C correspond to the timing of the variable in all subsequent tables.

debt are larger among low-quality managers, providing initial support for our hypothesis. We do not find consistent evidence when examining earnings quality based on the Dechow and Dichev measure, and explore this further in our multivariate analysis.

In Table 2 we find that managerial ability, measured with the MA-Score, is positively correlated with future earnings and negatively correlated with historical losses. Managerial ability is negatively correlated with restatements and errors in the provision for bad debt, and positively correlated with firm-specific earnings persistence. The Dechow and Dichev accruals quality measure, however, is negatively associated with the managerial ability, consistent with Francis et al. (2008) and our results in Panel B of Table 1.

IV. TEST DESIGN AND RESULTS

Earnings Restatements

Our first measure of earnings quality is earnings restatements, which are *ex post* evidence of erroneous reported earnings and thus have been used as a signal of poor earnings quality (Dechow et al. 2010). Although restatements can occur for reasons other than errors in accrual estimation, this earnings quality measure is the least reliant on an estimation procedure and, thus, provides a relatively unambiguous signal of earnings quality. Moreover, we expect restatements to be associated with errors in accrual estimation, as most restatements impact an accrual account (Palmrose and Shultz 2004). This assertion is supported by Plumlee and Yohn (2010), who find that a large number of restatements are a result of management's judgments and estimates.[14]

[14] Although some accrual errors will result in prospective adjustments, other mistakes in judgments and estimates will result in retroactive restatements (Plumlee and Yohn 2010). Plumlee and Yohn (2010) provide examples of how complexity and mistakes in judgments result in restatements. For example, firms have restated earnings when they chose to rely on their own estimate of fair value for a security whose last traded price was not necessarily indicative of fair value and *ex post* the SEC disagreed with the judgment made by management. As another example, firms have been required to restate earnings when subsequent information differs from management's expectations at the end of the fiscal year.

TABLE 2
Univariate Correlations

	Mgrl-Ability	*Historical Ret*	*Media Citations*	*Future Earnings*	*Loss%*	*Firm Size*	*Restate*	*F.S. EarnPer*	*BDE Error*	*AQ*	*Modified AQ_{LOSS}*	*Modified AQ_{SI}*	*Modified AQ_{CFO}*
MgrlAbility		**0.23**	**−0.10**	**0.27**	**−0.22**	0.00	**−0.02**	**0.04**	**−0.22**	**−0.02**	**0.03**	**0.03**	0.00
Historical Ret	**0.24**		−0.02	**0.41**	**−0.44**	**0.22**	**−0.03**	**0.13**	−0.05	**0.16**	**0.23**	**0.22**	**0.20**
Media Citations	**−0.09**	−0.01		**−0.10**	**0.02**	**0.52**	0.01	**−0.10**	−0.03	**0.11**	**0.07**	**0.08**	**0.10**
FutureEarnings	**0.08**	**0.26**	−0.02		**−0.56**	**0.29**	**−0.02**	**0.12**	**−0.17**	**0.23**	**0.33**	**0.32**	**0.29**
Loss%	**−0.21**	**−0.43**	0.00	**−0.33**		**−0.45**	**0.01**	**−0.07**	**0.32**	**−0.31**	**−0.42**	**−0.41**	**−0.37**
FirmSize	**−0.02**	**0.21**	**0.53**	**0.30**	**−0.45**		**0.06**	**0.02**	**−0.18**	**0.41**	**0.42**	**0.42**	**0.43**
Restate	**−0.02**	**−0.03**	0.01	−0.01	0.01	**0.06**		**−0.03**	−0.03	0.00	**−0.02**	**−0.01**	0.00
F.S. EarnPer	**0.04**	**0.13**	**−0.10**	**0.03**	**−0.06**	**0.02**	**−0.03**		**0.09**	**0.02**	**0.04**	**0.03**	**0.02**
BDE Error	**−0.16**	**−0.13**	0.11	**−0.13**	**0.32**	**−0.18**	0.00	**0.07**		0.02	0.04	0.04	0.06
AQ	**−0.02**	**0.17**	**0.11**	0.18	**−0.31**	**0.41**	0.00	**0.01**	0.02		**0.79**	**0.79**	**0.83**
Modified $AQ_{LOSS\%}$	**0.03**	**0.23**	**0.08**	**0.22**	**−0.41**	**0.42**	**−0.02**	**0.03**	−0.05	**0.78**		**0.92**	**0.87**
Modified $AQ_{SI\ Loss\%}$	**0.03**	**0.23**	**0.08**	**0.22**	**−0.41**	**0.43**	**−0.01**	**0.03**	−0.07	**0.79**	**0.92**		**0.87**
Modified $AQ_{NegCFO\%}$	0.00	**0.20**	**0.10**	**0.21**	**−0.37**	**0.44**	−0.01	**0.02**	−0.02	**0.83**	**0.87**	**0.87**	

Bold denotes significant correlation coefficients at the 10 percent alpha level.
This table reports Pearson correlation coefficients below the diagonal and Spearman correlation coefficients above the diagonal. We decile rank *MgrlAbility*, *Historical Ret*, *Media Citations*, *F.S. EarnPer*, *AQ*, and *AQ* modifications by industry-year.
See Table 1, Panel C for variable definitions.

Restate is an indicator variable that is equal to 1 if there is an announcement of a financial restatement in years t+1, 2, or 3. In our main analysis, we use the restatement data from Hennes et al. (2008) for restatements announced from 1997–2006 and from Audit Analytics for restatements announced from 2007–2010.[15] We then supplement our main analysis with those restatements identified as due to management's judgments and estimates per Plumlee and Yohn (2010), using the sample from their study.[16]

To determine whether managerial ability varies with earnings restatements, we estimate the following equation using a pooled logistic regression:

$$\begin{aligned} Restate_{t+1,t+3} = {} & \alpha_0 + \boldsymbol{\alpha_1 MgrlAbility_t} + \alpha_2 FirmSize_t + \alpha_3 SalesVolatilty_{t-4,t} \\ & + \alpha_4 CashFlowVolatility_{t-4,t} + \alpha_5 OperCycle_{t-4,t} + \alpha_6 Loss\%_{t-4,t} \\ & + \alpha_7 NationalAuditor_t + \alpha_8 \Delta SalesGrowth_t + \alpha_9 AbnormalReturn_t + \varepsilon_{t+1,t+3}. \end{aligned} \quad (2)$$

We include each of the control variables discussed above. Because our tests rely on panel data, standard errors may be correlated within years and across time by firm. Thus, unless otherwise noted, in this and all subsequent estimations we either cluster our standard errors by firm and year (Petersen 2009) or include firm fixed effects.

In Table 3 the first (second) estimation considers all restatements, and excludes (includes) firm fixed effects. As in our univariate analysis, we document a negative relation between managerial ability and restatements, supporting our hypothesis that more able managers are associated with higher quality earnings. The more efficient the manager, the less likely the firm is to restate ($\alpha_1 = -0.21$; $p < 0.05$). Given that the unconditional likelihood of having a restatement is 13 percent, untabulated results indicate that the marginal effect is economically significant at −2.4 percent.

In the second set of estimations in Table 3, we examine the restatements from 2003–2006 considered in Plumlee and Yohn (2010) and partition the sample between those restatements associated with management's judgments and estimates in column 3 and all other restatements in column 4. We find that only those restatements associated with management's judgments and estimates are associated with managerial ability.[17] These findings support our use of *Restate* as an earnings quality metric affected by managers' judgments and estimates. Again, however, a limitation of *Restate* as a measure of earnings quality is that many errors in judgments and accruals estimates will not result in a retrospective restatement, illustrating the importance of considering multiple earnings quality measures. We expect each of our remaining earnings quality measures to be lower in the presence of errors in judgments and estimates, even when these errors do not result in a restatement of previously issued financial statements.

Earnings Persistence

Our second measure of earnings quality is earnings persistence, which is frequently discussed as a measure of earnings quality (e.g., Dechow et al. 2010). We expect higher ability managers to choose better projects, have an improved understanding of risk, and manage the firm's operations more efficiently (by construction). Thus, we expect more able managers to have more persistent

---

15 We thank the authors for the GAO-based dataset, which is available at http://sbaleone.bus.miami.edu/.

16 We thank the authors for providing us both the restatement data and their coding of the restatements between those related to judgments and estimates and other restatements.

17 Other restatements include internal errors and manipulation. We do not estimate these specifications with firm fixed effects, which require variation in the dependent variable for estimation and, thus, result in a sample size of only 1,834 of the 10,568 observations.

American Accounting Association

TABLE 3

Restatements and Managerial Ability

$$Restate_{t+1,t+3} = \alpha_0 + \boldsymbol{\alpha_1 MgrlAbility_t} + \alpha_2 FirmSize_t + \alpha_3 SalesVolatility_{t-4,t} + \alpha_4 CashFlowVolatility_{t-4,t} + \alpha_5 OperCycle_{t-4,t} + \alpha_6 Loss\%_{t-4,t} + \alpha_7 NationalAuditor_t + \alpha_8 \Delta SalesGrowth_t + \alpha_9 AbnormalReturn_t + \varepsilon_{t+1,t+3}.$$

		Dependent Variable =			
	Pred.	***Restate***	***Restate***	***Restate Judgments***	***Restate Other***
MgrlAbility	–	−0.21**	−0.22**	−0.50**	−0.13
		−2.09	*−2.20*	*−2.33*	*−1.08*
FirmSize	+	0.13***	0.47***	−0.06	−0.02
		3.81	*11.39*	*1.26*	*0.35*
SalesVolatility	+	0.21*	−0.26	−0.05	0.83***
		1.62	*−1.58*	*−0.21*	*3.74*
CashFlowVolatility	+	0.07	−0.35	0.74	−0.33
		0.83	*−0.90*	*1.57*	*−0.88*
OperCycle	+	−0.08*	−0.11	−0.20***	0.10*
		−1.78	*−1.31*	*−2.86*	*1.83*
Loss%	+	0.32***	0.32***	0.03	0.29***
		3.47	*2.78*	*0.19*	*2.58*
NationalAuditor	–	−0.23	−0.09	0.42	0.24
		−1.58	*−1.04*	*1.60*	*0.98*
ΔSalesGrowth	?	0.00	0.00	0.00	0.02
		0.53	*0.26*	*0.00*	*0.94*
AbnormalReturn	?	0.05	0.04	0.07	0.02
		1.00	*1.54*	*0.92*	*0.20*
Restatement Obs.		4,453	4,453	1,104	1,393
Total Observations		33,035	12,182	10,568	10,568
Pseudo R^2		1.00%	NA	1.14%	0.85%
Firm Fixed Effects		Excluded	Included	Excluded	Excluded

*, **, *** Denote a two-tailed p-value of less than 0.10, 0.05, and 0.01, respectively.
This table reports the results from the logistic regression of earnings restatements on managerial ability and controls for innate firm characteristics. Z-statistics are presented in italics below the coefficients and are based on standard errors that are clustered by firm and year for specifications excluding firm fixed effects. We decile rank *MgrlAbility* by industry-year. Intercept is included, but not tabulated.
See Table 1, Panel C for variable definitions.

earnings, and expect this effect to influence both accruals and cash flows, both of which are designed to reflect the underlying economics of the firm.[18]

Prior research has shown that accruals tend to have a lower persistence than cash flows and one reason for this is that they contain more uncertainty, thereby requiring managerial estimation (e.g.,

[18] It is possible for managers to artificially smooth earnings and thus appear to have higher earnings persistence. This artificial persistence should not extend, however, to the breakdown of accruals and cash flows.

Richardson et al. 2005). Thus, in addition to more persistent accruals and cash flows related to operations, we expect an incremental effect of managerial ability on the persistence of the accrual component of earnings. In sum, we have two expectations related to our hypothesis that higher ability managers report higher quality earnings: (1) earnings reported by higher ability managers are more persistent than earnings reported by lower ability managers due to both superior operations and superior accrual estimation, and (2) higher ability managers' impact on accruals exceeds their impact on cash flows because the former reflects both operational efficiency *and* superior accrual estimation. We examine earnings persistence using the following model:

$$\begin{aligned} Earnings_{t+1,t+n} = {} & \alpha_0 + \alpha_1 Earnings_t + \boldsymbol{\alpha_2 Earnings_t \times MgrlAbility_t} + \alpha_3 MgrlAbility_t \\ & + \alpha_4 FirmSize_t + \alpha_5 SalesVolatility_{t-4,t} + \alpha_6 CashFlowVolatility_{t-4,t} \\ & + \alpha_7 OperCycle_{t-4,t} + \alpha_8 Loss\%_{t-4,t} + \alpha_9 NationalAuditor_t \\ & + \alpha_{10} \Delta SalesGrowth_t + \alpha_{11} AbnormalReturn_t + \varepsilon_{t+1,t+n}. \end{aligned} \tag{3}$$

We calculate earnings as earnings before extraordinary items (Xpressfeed [hereafter "XFN"] variable = IBC) scaled by average total assets (XFN = AT) and then separate earnings into accruals and cash flow components. Because earnings persistence is not desirable for loss firms, we estimate profit and loss firms separately and only tabulate results for profit firms.[19]

In Table 4 we examine the impact of managerial ability on earnings persistence and then more formally examine our hypothesis by investigating the relative impact of ability on the persistence of accruals and cash flows. In the first column of estimates, which control for firm fixed effects, the base persistence is 0.30 and is increasing with managerial ability. Earnings persistence is expected to increase from 0.30 to 0.66 (0.30 + 0.36) for firms with positive earnings when moving from the lowest to the highest decile of managerial ability. Although not tabulated, we find similar results when firm fixed effects are excluded.

When we partition earnings into accruals and cash flows, managerial ability increases the persistence of both components, but increases the accrual component more than the cash flow component. The accruals reported by positive earnings firms have a base persistence of 0.36 when firm fixed effects are excluded. The incremental coefficient on accruals for firms with higher ability managers is 0.42 ($p < 0.01$). In comparison, the base persistence of cash flows is 0.69, and the incremental coefficient on cash flows for firms with higher ability managers is 0.26 ($p < 0.01$), and this incremental effect is statistically smaller than that of accruals ($p < 0.05$). We find similar results when we include firm fixed effects in the model. These findings support our hypothesis that higher quality managers are better able to estimate accruals, resulting in higher earnings quality, and also support our prediction that higher quality managers operate their businesses more effectively.

We also replace one-year-forward earnings with average earnings from year t+1 to t+3 in order to reduce the impact of economic shocks occurring in any particular year. Results are similar, although the interaction between managerial ability and accruals becomes insignificant when firm fixed effects are included.

McNichols and Wilson (1988) Error in the Provision for Bad Debt

Our third measure of earnings quality is the provision for bad debt, modeled in McNichols and Wilson (1988), as follows:

[19] For this test, in order to continue to present the dependent variable in year t+1, we include earnings in year t, the same period in which managerial ability is estimated. Results are not sensitive to this timing choice. For example, untabulated results indicate that conclusions remain unchanged if we continue to measure managerial ability at time t, but consider how much of year t+1 earnings persists into year t+2.

TABLE 4

Earnings Persistence and Managerial Ability

$$Earnings_{t+1,t+n} = \alpha_0 + \alpha_1 Earnings_t + \boldsymbol{\alpha_2 Earnings_t \times MgrlAbility_t} + \alpha_3 MgrlAbility_t + \alpha_4 FirmSize_t + \alpha_5 SalesVolatility_{t-4,t} + \alpha_6 CashFlowVolatility_{t-4,t} + \alpha_7 OperCycle_{t-4,t} + \alpha_8 Loss\%_{t-4,t} + \alpha_9 NationalAuditor_t + \alpha_{10} \Delta SalesGrowth_t + \alpha_{11} AbnormalReturn_t + \varepsilon_{t+1,t+n}.$$

		Dependent Variable =					
	Pred.	*Future* $Earnings_{t+1}$			*Average Future* $Earnings_{t+1,t+3}$		
Earnings	+	0.30***			0.25***		
		9.01			*6.81*		
Earnings × MgrlAbility	+	0.36***			0.09*		
		7.90			*1.74*		
Accruals	+		0.36***	0.23***		0.31***	0.24***
			5.18	*6.64*		*4.99*	*6.25*
Accruals × MgrlAbility	+		0.42***	0.35***		0.41***	0.03
			6.55	*7.14*		*4.71*	*0.51*
CFO	+		0.69***	0.45***		0.62***	0.30***
			12.63	*12.36*		*7.52*	*7.37*
CFO × MgrlAbility	+		0.26***	0.26***		0.28***	0.08
			5.01	*5.46*		*2.84*	*1.57*
MgrlAbility	?	−0.01**	0.00	0.00	0.00	0.00	0.00
		−2.01	*0.66*	*0.49*	*0.01*	*0.62*	*0.13*
Total Observations		33,735	33,735	33,735	28,016	28,016	28,016
R^2		2.07%	15.96%	3.11%	0.07%	10.20%	0.01%
Test (F-statistic): *Accruals × MgrlAbility = CFO × MgrlAbility*			5.62**	7.04***		2.66*	2.41
Firm Fixed Effects		Included	Excluded	Included	Included	Excluded	Included

*, **, *** Denote a two-tailed p-value of less than 0.10, 0.05, and 0.01, respectively.

(continued on next page)

TABLE 4 (continued)

This table presents the OLS regression results investigating the relation between managerial ability and earnings persistence for firms with positive earnings in year *t*. t-statistics are presented in italics below the coefficients and are based on standard errors that are clustered by firm and year in specifications excluding firm fixed effects. We decile rank *MgrlAbility* by industry-year. Our main set of control variables (*FirmSize*, *SalesVolatility*, *CashFlowVolatility*, *OperCycle*, *Loss%*, *ΔSalesGrowth*, *AbnormalReturn*, and *NationalAuditor*) are included in the model. For succinctness, however, results for the control variables are not tabulated. Intercept is included, but not tabulated.
See Table 1, Panel C for variable definitions.

$$Bad\ Debt\ Expense_t = \alpha_0 + \alpha_1 Allowance\ for\ Doubtful\ Accounts_{t-1} + \alpha_2 Write\text{-}Offs_t + \alpha_3 Write\text{-}Offs_{t+1} + \varphi_t, \tag{4}$$

where *Bad Debt Expense* and *Write-Offs* are hand-collected from the firm's SEC filings and *Allowance for Doubtful Accounts* is available from Xpressfeed (XFN = RECD). All variables are deflated by sales in year *t*. Implicit in our use of this model is a balance sheet perspective to estimating bad debt, adherence to GAAP, and perfect foresight of future write-offs.

The error (φ_t) has two components: a discretionary "earnings management" component and a forecast error component (McNichols and Wilson 1988). If managers' estimates are unbiased, on average, then errors in the bad debt accrual will vary with forecast accuracy, and we expect this error (φ_t) to decrease with managerial ability. Jackson and Liu (2010), however, present evidence that the mean earnings management portion of φ_t is positive; specifically, they find that managers tend to overstate the allowance for doubtful accounts that they can later reverse into income. Under the assumption that all managers engage in the same degree of earnings management, the variation in the bad debt error will be driven by variation in accrual-estimation quality, and thus we continue to expect the error to decrease with managerial ability.[20]

Because the data for this analysis must be hand-collected from SEC filings, we limit the analysis to firms with managers in the top or bottom quintiles of managerial ability and, following McNichols and Wilson (1988), to three accounts-receivable-intensive industries: (1) printing and publishing, (2) nondurable wholesale goods, and (3) business services, to estimate:

$$\begin{aligned} BDE\ Error_{t+1} = {} & \alpha_0 + \boldsymbol{\alpha_1 HighAbilityIndicator_t} + \alpha_2 FirmSize_t + \alpha_3 SalesVolatility_{t-4,t} \\ & + \alpha_4 CashFlowVolatility_{t-4,t} + \alpha_5 OperCycle_{t-4,t} + \alpha_6 Loss\%_{t-4,t} \\ & + \alpha_7 NationalAuditor_t + \alpha_8 \Delta SalesGrowth_t + \alpha_9 AbnormalReturn_t + \varepsilon_{t+1}, \end{aligned} \tag{5}$$

where *BDE Error* is the absolute value of the residual from Equation (4), and *HighAbilityIndicator* is an indicator variable that is equal to 1 (0) if the managerial ability score in year *t* is in the top (bottom) quintile relative to industry-year peers. A negative coefficient on *HighAbilityIndicator* is consistent with more able managers forming better estimates of bad debt provisions. Results are presented in Table 5. In support of our hypothesis, α_1 is -0.01 ($p < 0.01$), consistent with managers with higher ability scores producing higher quality bad debt provisions.[21]

The Dechow and Dichev Measure of Accruals Quality

Our final measure of earnings quality follows Dechow and Dichev (2002), who posit that high-quality accruals are eventually realized as cash flows. Incorrectly estimated accruals are less likely to be realized as cash flows. We hypothesize that the better managers know their business, the less likely they are to have erroneous accruals. We determine how well a firm's accruals map into cash flows by estimating the following regression by industry (Fama and French 1997) and year:[22]

---

[20] Alternatively, if higher ability managers engage in more earnings management, then our tests are conservative. Only if higher ability managers engage in less earnings management would the findings of Jackson and Liu (2010) weaken the basis for our conclusions. In this setting, the negative relation between managerial ability and the bad debt error (φ_t) could result from better estimations, less earnings management, or both. In this latter case, the bad debt error analysis is a weak test of the relation between managerial ability and estimation quality, but continues to speak to the broader conclusions regarding earnings quality.

[21] Because our sample selection procedure for this analysis results in a small sample of firm-years that do not necessarily contain the same firm over multiple years (i.e., we do not have panel data), we do not cluster standard errors by firm and year, nor do we estimate a firm fixed effects specification. In untabulated results, however, we continue to observe a negative and significant relation between managerial ability and *BDE Error* when we cluster standard errors by year or include year fixed effects.

[22] We delete those observations for which the industry group has less than 20 observations in any given year.

TABLE 5

Errors in the Allowance for Bad Debt and Managerial Ability

$$BDE\ Error_{t+1} = \alpha_0 + \boldsymbol{\alpha_1 HighAbilityIndicator_t} + \alpha_2 FirmSize_t + \alpha_3 SalesVolatilty_{t-4,t} + \alpha_4 CashFlowVolatility_{t-4,t} + \alpha_5 OperCycle_{t-4,t} + \alpha_6 Loss\%_{t-4,t} + \alpha_7 NationalAuditor_t + \alpha_8 \Delta SalesGrowth_t + \alpha_9 AbnormalReturn_t + \varepsilon_{t+1}.$$

	Pred.	Dependent Variable = *BDE Error*
HighAbilityIndicator	–	−0.01***
		−2.77
FirmSize	+	0.00
		1.13
SalesVolatility	+	0.00
		1.25
CashFlowVolatility	+	0.02*
		1.65
OperCycle	+	0.01
		1.74
Loss%	+	0.01***
		4.24
NationalAuditor	–	0.00
		0.60
ΔSalesGrowth	?	0.00
		1.32
AbnormalReturn	?	0.00
		1.12
Total Observations		838
R^2		9.00%

*, *** Denote a two-tailed p-value of less than 0.10 and 0.01, respectively.
This table presents the OLS regression results investigating the relation between managerial ability and errors in the bad debt provision. t-statistics are presented in italics below the coefficients (we do not cluster by firm and year or include firm fixed effects because we do not have a sufficient number of same-firm observations). *BDE Error* is the absolute value of the residual from Equation (5), and *HighAbilityIndicator* is an indicator variable that is equal to 1 (0) if the managerial ability score in year *t* is in the top (bottom) quintile relative to industry-year peers. Requisite information for this test requires hand-collection from SEC filings. Thus, we limit the analysis to firms in three industries (following McNichols and Wilson [1988]) where accounts receivable (relative to assets) and bad debt expense (relative to earnings) are large: (1) printing and publishing; (2) nondurable wholesale goods; and (3) business services. We consider only those firm-year observations where managerial ability falls among the highest and lowest quintile relative to industry-year peers. Intercept is included, but not tabulated.
See Table 1, Panel C for variable definitions.

$$\Delta WC_t = \alpha_0 + \alpha_1 CFO_{t-1} + \alpha_2 CFO_t + \alpha_3 CFO_{t+1} + \alpha_4 \Delta REV_t + \alpha_5 PPE_t + \varepsilon_t. \tag{6}$$

The residual from the regression measures the extent to which current accruals map into past, present, or future cash flows, with smaller absolute residuals indicating superior mapping. Following McNichols (2002), we include the current-year change in sales (ΔREV) and the current-

year level of property, plant, and equipment (*PPE*) in Equation (6); variable definitions are in Table 1.

Following prior research, we take the standard deviation of the residual over a rolling four-year period, and multiply this standard deviation by −1 so that the variable is increasing with earnings quality. Thus, $AQ_t = -1 \times$ Standard Deviation (ε_{t+1}, ε_{t+2}, ε_{t+3}, ε_{t+4}).[23] To maintain consistency with the MA-Score, we create decile ranks of our earnings quality variable by year and industry.

To examine our hypothesis using this measure of earnings quality, we estimate the following regression:

$$\begin{aligned} AQ_{t+1,t+4} = {} & \alpha_0 + \boldsymbol{\alpha_1 MgrlAbility_t} + \alpha_2 FirmSize_t + \alpha_3 SalesVolatility_{t-4,t} \\ & + \alpha_4 CashFlowVolatility_{t-4,t} + \alpha_5 OperCycle_{t-4,t} + \alpha_6 Loss\%_{t-4,t} \\ & + \alpha_7 NationalAuditor_t + \alpha_8 \Delta SalesGrowth_t + \alpha_9 AbnormalReturn_t + \varepsilon_{t+1,t+4}. \end{aligned} \quad (7)$$

The results in Table 6 indicate that, counter to our expectations, but consistent with Francis et al. (2008), earnings quality is decreasing in managerial ability when firm fixed effects are excluded, although managerial ability is insignificant when firm fixed effects are included. Prior work notes that some accruals are simply more difficult to estimate, and this estimation difficulty is associated with the firm's operating environment (Dechow and Dichev 2002; McNichols 2002; LaFond 2008). Thus, Francis et al. (2008) conclude that better managers are hired to manage more challenging firms, thereby providing an explanation for this counter-intuitive finding.

Dechow et al. (2010) and others note that most model-based accruals quality measures contain a large firm-specific component. The Dechow and Dichev model constrains the coefficient on cash flows to be the same across observations (e.g., Ball and Shivakumar 2006; Wysocki 2009) and, thus, the resulting measure of accruals quality contains measurement error that is systematically associated with firm characteristics, such as volatility. Although we control for innate factors known to affect accruals quality, it is possible that the effect of these innate factors varies across firms and, thus, the controls are inadequate.

Ball and Shivakumar (2006) modify the Dechow and Dichev model by allowing the coefficient on current-period cash flows to vary among observations with negative cash from operations, their proxy for economic losses, and document a large degree of variation.[24] To illustrate the effect of this modification, in their Table 3, Panel A, the pooled coefficient on current-period operating cash flows is −0.57 among firms with positive operating cash flows, while it is −0.12 (−0.57 + 0.45) among firms with negative operating cash flows. Absent this type of modification, the coefficient on cash flows is forced to be the same across firms, and differences in the expected relation between accruals and cash flows are relegated to the residual term, affecting our measure of earnings quality. Again, it is unlikely that standard control variables, which are also constrained to have the same effect across firms, will be sufficient to ameliorate this effect. Thus, it is possible that better

---

23 Because managers may not necessarily be in place for the full aggregation period, we also consider a one-year accrual error by considering the absolute value of the residual from Equation (6) directly. Dechow and Dichev (2002, footnote 6) also estimate a one-year accrual error. The greater the residual, in absolute terms, the poorer the accruals quality. Results are quantitatively and qualitatively similar using this modification; thus, we focus on the traditional measure in our main analyses for succinctness. In our additional analyses, however, where we track managers across firms, we use the annual residual to better match accruals quality with specific managers.

24 Ball and Shivakumar (2006) highlight that the matching effect of accruals is expected to lead to a negative correlation between accruals and cash flows (e.g., Dechow 1994). In contrast, they note that opposite to the noise-reducing role of operating accruals, the gain and loss recognition role of accruals is a source of positive correlation between accruals and current-period operating cash flow (Ball and Shivakumar 2006, 212–213). They then examine several cash flow and market-return-based measures of economic losses. Note that in Panel A of Table 7, we document a positive association between accruals and cash flows for the quintile of firms with the most losses.

TABLE 6

Accruals Quality and Managerial Ability

$$AQ_{t+1,t+4} = \alpha_0 + \boldsymbol{\alpha_1 MgrlAbility_t} + \alpha_2 FirmSize_t + \alpha_3 SalesVolatility_{t-4,t} + \alpha_4 CashFlowVolatility_{t-4,t} + \alpha_5 OperCycle_{t-4,t} + \alpha_6 Loss\%_{t-4,t} + \alpha_7 NationalAuditor_t + \alpha_8 \Delta SalesGrowth_t + \alpha_9 AbnormalReturn_t + \varepsilon_{t+1,t+4}.$$

		Dependent Variable =		
	Pred.		*AQ*	
MgrlAbility	+	−0.03***		0.00
		−3.39		*0.52*
FirmSize	+	0.04***		0.02***
		17.11		*7.63*
SalesVolatility	−	−0.14***		0.05***
		−8.53		*4.48*
CashFlowVolatility	−	−0.42***		0.15***
		−8.08		*4.94*
OperCycle	−	0.00		0.00
		0.04		*0.13*
Loss%	−	−0.11***		−0.04***
		−8.17		*−4.56*
NationalAuditor	+	0.02***		0.02***
		2.66		*3.28*
ΔSalesGrowth	?	−0.01*		0.00
		−1.64		*0.15*
AbnormalReturn	?	0.01***		0.00
		4.57		*0.83*
Total Observations			31,957	
R^2		18.93%		9.83%
Firm Fixed Effects		Excluded		Included

*, *** Denote a two-tailed p-value of less than 0.10 and 0.01, respectively.
This table reports the results from the OLS regression of accruals quality on managerial ability and controls for innate firm characteristics. t-statistics are presented in italics below the coefficients and are based on standard errors that are clustered by firm and year in specifications excluding firm fixed effects. We decile rank *MgrlAbility* and *AQ*, by industry-year. Intercept is included, but not tabulated.
See Table 1, Panel C for variable definitions.

managers can improve the estimation quality of accruals, but the resulting accruals quality still falls below that of a firm with closer-to-average associations between accruals and cash flows.

Although Ball and Shivakumar (2006) focus on the effect of economic losses, Wysocki (2009) notes that a number of different innate factors affect the relation between accruals and cash. Thus, rather than focusing solely on differences in current-period cash flows, we investigate a broader range of firm characteristics. In Panel A of Table 7 we examine the correlation between working capital accruals and current-period cash from operations for each quintile rank of the "innate" firm characteristics from Francis et al. (2004), such as firm size, losses, and operating volatility, as well

TABLE 7

Accruals, Cash Flows, and Firm Characteristics

Panel A: Correlation between Working Capital Accruals and Current-Period Cash Flows within Innate Quintiles

	Innate Characteristic Quintiles =					
Innate Characteristic	**1**	**2**	**3**	**4**	**5**	**Difference**
FirmSize						
corr(ΔWC_t, CFO_t)	0.09***	−0.22***	−0.28***	−0.33***	−0.36***	−0.45
SalesVolatility						
corr(ΔWC_t, CFO_t)	−0.05***	−0.08***	−0.09***	−0.09***	0.06***	0.11
CashFlowVolatility						
corr(ΔWC_t, CFO_t)	−0.11***	−0.13***	−0.17***	−0.15***	0.03***	0.14
OperCycle						
corr(ΔWC_t, CFO_t)	0.00	−0.09***	−0.12***	−0.09***	0.07***	0.07
Loss%						
corr(ΔWC_t, CFO_t)	−0.57***	−0.50***	−0.21***	−0.07***	0.03***	0.60
Loss% Before SI						
corr(ΔWC_t, CFO_t)	−0.52***	−0.50***	−0.19***	−0.08***	0.04**	0.56
Negative CFO%						
corr(ΔWC_t, CFO_t)	−0.32***	−0.30***	−0.28***	−0.06***	0.11***	0.42

Panel B: The Mean Coefficients from Regressions of Working Capital Accruals on Cash Flows, Change in Revenue, and PPE Estimated by Industry and *Loss%* Quintile

	Loss% Quintile =					
	1	**2**	**3**	**4**	**5**	**Traditional: Estimated by Industry-Year**
	Fewer Losses			**More Losses**		
CFO_{t-1}	0.15*	0.22***	0.06	0.21***	0.18***	0.24***
CFO	−0.56***	−0.58***	−0.44***	−0.42***	−0.30***	−0.38***
CFO_{t+1}	0.16***	0.11***	0.02	0.12***	0.12***	0.15***
ΔRev	0.04	0.10***	0.10***	0.10***	0.10***	0.09***
PPE	0.02	0.00	−0.01	−0.03***	−0.01	−0.02***
Intercept	0.03**	0.04***	0.03***	0.02***	0.00	0.01***
R^2	66%	58%	52%	44%	37%	45%

(continued on next page)

as losses before special items (Dechow and Ge 2006) and negative cash flows (Ball and Shivakumar 2006).[25]

Of these innate firm characteristics, the proportion of losses from year t–4 to year t has the starkest difference in correlations across quintiles. In Panel A of Table 7, the correlation between operating cash flows and accruals varies from −0.57 among firms with the fewest losses to 0.03 among firms with the most losses. The range across quintiles also varies widely, with ranges of 0.56

[25] We examine five years of cash flows, analogous to our loss proportion measure.

TABLE 7 (continued)

Panel C: The Mean Coefficients from Regressions of Working Capital Accruals on Cash Flows, Change in Revenue, and *PPE* Estimated by Industry and *Loss%* Adjusted for *SI* Quintile

	Loss% before *SI* Quintile =					
	1	2	3	4	5	Traditional: Estimated by Industry-Year
	Fewer Losses				More Losses	
CFO_{t-1}	0.22***	0.19***	0.13***	0.18***	0.16***	0.24***
CFO	−0.63***	−0.55***	−0.50***	−0.43***	−0.22***	−0.38***
CFO_{t+1}	0.15***	0.10***	0.08***	0.13***	0.11***	0.15***
ΔRev	0.08***	0.06***	0.11***	0.10***	0.09***	0.09***
PPE	0.03**	−0.01	0.00	−0.03***	0.00	−0.02***
Intercept	0.03***	0.04***	0.02***	0.02***	−0.01	0.01***
R^2	62%	55%	53%	45%	36%	45%

Panel D: The Mean Coefficients from Regressions of Working Capital Accruals on Cash Flows, Change in Revenue and *PPE* Estimated by Industry and *Negative CFO%* Quintile

	Negative CFO% Quintile =					
	1	2	3	4	5	Traditional: Estimated by Industry-Year
	Positive *CFO*				Negative *CFO*	
CFO_{t-1}	0.18***	0.22***	0.19***	0.16***	0.21***	0.24***
CFO	−0.53***	−0.49***	−0.53***	−0.45***	−0.25***	−0.38***
CFO_{t+1}	0.18***	0.13***	0.17***	0.12***	0.14***	0.15***
ΔRev	−0.09	0.08***	0.10***	0.11***	0.09***	0.09***
PPE	0.01	0.00	−0.01	−0.02*	−0.03*	−0.02***
Intercept	0.02***	0.02***	0.02*	0.02***	0.03***	0.01***
R^2	50%	43%	54%	49%	37%	45%

*, **, *** Denote a Fama-MacBeth p-value of less than 0.10, 0.05, and 0.01, respectively.
Year indicator variables are included in models estimated by industry and innate firm characteristic rank, but coefficients are not tabulated.
See Table 1, Panel C for variable definitions.

for the proportion of losses before special items, −0.45 for firm size, and 0.42 for the proportion of negative cash flows. In subsequent analyses discussed below, we exclude firm size as a partitioning variable for two reasons. First, the variation in firm size is driven almost entirely by the smallest size quintile, while the other partitioning variables have more monotonic changes in the association between accruals and cash flows. Second, in our setting, we are concerned with systematic biases in the accruals quality measure that are associated with managerial ability, but from Table 2, managerial ability and firm size are not associated.[26]

For three partitioning variables (*Loss%*, *Loss% Before SI*, and *Negative CFO%*), we re-estimate Equation (6) and allow the coefficients on each of the variables to vary by each of these quintiles, i.e., to vary with the magnitude of the negative relation between accruals and cash flows

[26] Results are not affected by our treatment of firm size. Untabulated results are similar if we exclude the smallest size quintile from our sample, and are also similar *within* the smallest size quintile.

as measured by the quintile rank of the firm's fundamentals. Thus, we estimate a modification of Equation (6):

$$\Delta WC_t = \alpha_0 + \alpha_1 CFO_{t-1} + \alpha_2 CFO_t + \alpha_3 CFO_{t+1} + \alpha_4 \Delta REV_t + \alpha_5 PPE_t + Year\ Indicators + \varepsilon_t. \tag{8}$$

We estimate Equation (8) by the various quintiles and industry, and include year fixed effects. As illustrated in Panel B of Table 7, allowing the coefficients on the independent variables to vary with *Loss%* results in large differences across the estimations. For example, the mean R^2 falls from 66 percent in the group with the fewest losses to 37 percent in the group with the most losses. With particular reference to the preceding measurement error, the coefficient on current-period cash flows varies from a mean of –0.56 to a mean of –0.30, while the traditional estimation procedure restricts the mean coefficient to be approximately –0.38.[27]

Results are similar for the proportion of losses before special items, with the coefficient on current-year operating cash flows ranging from –0.63 to –0.22. Finally, the coefficients on current-period operating cash flows range from –0.53 to –0.25 across negative cash flows quintiles.

We next explore the extent to which these modifications are associated with (1) other measures of earnings quality and (2) managerial ability. The $ModifiedAQ_{Loss\%}$ accruals quality measure appears to be more correlated with our other measures of earnings quality than the traditional Dechow and Dichev accruals quality measure. For example, while *AQ* is not correlated with restatements, $ModifiedAQ_{Loss\%}$ is negatively correlated with restatements, as we would expect. Of particular interest for our study, $ModifiedAQ_{Loss\%}$ is positively associated with managerial ability, in contrast to the traditional model of accruals quality that exhibits a negative association with managerial ability.

The correlations for $ModifiedAQ_{SI\_Loss\%}$ are similar to those of $ModifiedAQ_{Loss\%}$, which is not surprising as these two measures are correlated at 0.92. Our third modification, $ModifiedAQ_{Neg\_CFO\%}$, does not appear to be as strongly associated with the other measures of earnings quality. Like the traditional measure, it is not associated with restatements, and the association with earnings persistence is the same as the traditional measure. Finally, the association with managerial ability is not significant.

We explore the relation between the modified measures of accruals quality and managerial ability more formally in Table 8.[28] In five of the six estimations, we find that managerial ability is positively associated with accruals quality, consistent with the relations with restatements, earnings persistence, and bad debt expense. Thus, it appears that the counter-intuitive result in our Table 6 and in Francis et al. (2008) reflects measurement error in the Dechow and Dichev accruals quality measure. In sum, we find that higher ability managers report lower accruals quality errors relative to industry-peers with similar operating environments.

V. ADDITIONAL ANALYSES

In this section we discuss three sets of results aimed at corroborating the conclusions from our main analyses. First, we explore the robustness of our results to alternative measures of managerial ability. Second, we investigate whether earnings quality increases when a new CFO of superior

[27] All coefficients are the mean coefficient across estimations; for the traditional model, regressions are estimated by industry and year; for the modified model, regressions are estimated by industry and *Loss%* quintile, and we include year fixed effects. Estimating the modified model by industry, *Loss%* quintile, and year results in a severe loss of observations.

[28] Because we condition on losses in our estimation of two of the modifications, we do not include losses as a control variable.

ability is hired. Third, we investigate the sensitivity of our results to the inclusion of additional control variables in our models.

Alternative Managerial Ability Measures

Our main managerial ability measure, the MA-Score, is the managerial efficiency metric developed by Demerjian et al. (2012). In this section we investigate historical returns, following Fee and Hadlock (2003), and media citations, following Francis et al. (2008); definitions are provided in Panel C of Table 1.[29] We begin by examining the correlations among the variables (see Table 2). Managerial ability and historical returns are correlated at 0.23, consistent with these two variables measuring different aspects of "ability," while there is a *negative* correlation between media citations and ability and no relation between media citations and historical returns. Consistent with historical returns and media citations containing a large firm component, they are correlated with firm size at 0.22 and 0.52, respectively, based on Spearman correlation coefficients, while managerial ability is not correlated with firm size. Interestingly, both ability and historical returns are negatively correlated with losses and positively correlated with future earnings; however, these associations do not extend to media citations. Generally, the correlations suggest that historical returns and the MA-Score have the expected associations with perceptions of managerial ability.

To examine the association between earnings quality and managerial ability, we consider a composite measure of total earnings quality, which is the sum of (1) the rank of one modified accruals quality measure ($ModifiedAQ_{Loss\%}$), (2) the rank of firm-specific earnings persistence, and (3) $-1 \times Restate$.[30] Thus, the variable ranges from a low of -1 to a high of 2. In Table 9, we consider the associations between total earnings quality and each of the three ability measures.[31] For the MA-Score, more able managers have higher total earnings quality, as was the case with the individual components. Next, we find that the greater the historical returns, the greater total earnings quality, consistent with our hypothesis. Untabulated results indicate that our inferences remain unchanged if we use historical returns as an alternate ability measure in each of our tests in the paper. However, as previously noted, the MA-Score allows us to better separate the manager from the firm.

Media citations, however, are negatively associated with earnings quality. Perhaps media interest increases as restatements and other negative earnings-related items occur. One difficulty in comparing results across measures is that media citations are available only for a small subset of our sample; thus, we next consider all three measures simultaneously for the reduced sample. The coefficient on *Media Citations* remains negative and significant, while the coefficients on *Historical Ret* and *MgrlAbility* remain positive and significant in the reduced sample.

CFO-Specific Analysis

Although our results are similar using historical returns, both the MA-Score and historical returns likely contain a firm-specific element. To better control for the effect of the firm, we next consider how earnings quality changes across different CFO regimes. We first explore whether there are CFO fixed effects, following Dyreng et al. (2010) and others. We then explore whether the CFO's ability at his/her initial firm is associated with the accruals quality at the subsequent firm. Together, these analyses provide evidence on the CFO-specific effect on accruals quality.

---

[29] Following our examination of these alternate measures, we also examine manager fixed effects.

[30] We exclude the bad debt expense error as it severely limits the available observations.

[31] Untabulated results are similar when we consider the individual components of total earnings quality, with the exception that the relation between historical returns and firm-specific earnings persistence is not significant. As with the results reported in Table 8, we exclude *Loss%* as a control variable because it is used to calculate $ModifiedAQ_{Loss\%}$. Results are similar if we include *Loss%* as an additional control variable.

TABLE 8

Modified Accruals Quality and Managerial Ability

$$\begin{aligned} \textit{Modified Accruals Quality}_{t+1,t+4} = {} & \alpha_0 + \alpha_1 \boldsymbol{MgrlAbility}_t + \alpha_2 FirmSize_t + \alpha_3 SalesVolatility_{t-4,t} \\ & + \alpha_4 CashFlowVolatility_{t-4,t} + \alpha_5 OperCycle_{t-4,t} \\ & + \alpha_6 NationalAuditor_t + \alpha_7 \Delta SalesGrowth_t \\ & + \alpha_8 AbnormalReturn_t + \varepsilon_{t+1,t+4}. \end{aligned}$$

		Dependent Variable =					
	Pred.	***Modified AQ$_{Loss\%}$***		***Modified AQ$_{SI_Loss\%}$***		***Modified AQ$_{NegCFO\%}$***	
MgrlAbility	+	0.06***	0.03***	0.05***	0.02***	0.03***	0.01
		6.53	*4.05*	*5.55*	*3.70*	*3.26*	*1.28*
FirmSize	+	0.04***	0.02***	0.04***	0.02***	0.04***	0.02***
		23.57	*8.93*	*22.44*	*9.14*	*21.69*	*9.99*
SalesVolatility	–	–0.16***	0.05***	–0.15***	0.06***	–0.15***	0.07***
		–9.35	*4.74*	*–9.23*	*5.71*	*–8.02*	*6.06*
CashFlowVolatility	–	–0.61***	0.11***	–0.59***	0.11***	–0.59***	0.14***
		–10.95	*3.39*	*–10.24*	*3.71*	*–9.85*	*4.45*
OperCycle	–	0.00	–0.01	0.00	0.00	–0.01	–0.01*
		1.14	*–0.54*	*–0.06*	*–0.96*	*–1.17*	*–1.85*
NationalAuditor	+	0.01	0.02**	0.01	0.02***	0.02**	0.03***
		1.44	*2.88*	*1.42*	*2.63*	*2.22*	*4.58*
ΔSalesGrowth	?	0.00	0.00	0.00	0.00	0.00	0.01
		–0.85	*–0.06*	*–0.92*	*–0.33*	*–0.94*	*0.43*
AbnormalReturn	?	0.02***	0.003*	0.02***	0.01**	0.02***	0.01**
		4.42	*1.76*	*4.85*	*2.25*	*4.03*	*2.26*
Total Observations		28,119		28,119		28,119	
R^2		19.06%	9.35%	19.14%	8.72%	19.98%	9.14%
Firm Fixed Effects		Excluded	Included	Excluded	Included	Excluded	Included

*, **, *** Denote a two-tailed p-value of less than 0.10, 0.05, and 0.01, respectively.
This table reports the results from the OLS regression of modified accruals quality on managerial ability and controls for innate firm characteristics. t-statistics are presented in italics below the coefficients and are based on standard errors that are clustered by firm and year in specifications excluding firm fixed effects. We decile rank *MgrlAbility* and the *AQ* modifications, by industry-year. Intercept is included, but not tabulated.
See Table 1, Panel C for variable definitions.

We identify 195 CFOs who switch employment among firms covered by ExecuComp, which identifies the CFOs, and are able to estimate CFO fixed effects for 62 of these executives across 158 firms. In untabulated results, we estimate the following:

$$TotalEarnQuality2_{i,t} = \alpha_0 + \Sigma_t \alpha_t YEAR_t + \Sigma_i \alpha_i FIRM_t + \Sigma_m \alpha_m MANAGER_m + \varepsilon_{i,t}. \tag{9}$$

TotalEarnQuality2 is defined as *TotalEarnQuality* except that we measure accruals quality as the decile rank within industry year of the *annual* accrual error from the modified accruals model in Equation (8), estimated by historical loss proportion ($ModifiedAQ_{Loss\%Ann}$).[32] This approach better

[32] Both Jones et al. (2008) and Dechow and Dichev (2002) use the absolute value of the annual residual from the Dechow and Dichev model as a measure of accruals quality.

TABLE 9

Earnings Quality and Alternative Proxies for Managerial Ability

$$\begin{aligned} TotalEarnQuality_{t+1,t+4} = {} & \alpha_0 + \boldsymbol{\alpha_1 MgrlAbility_t} + \alpha_2 FirmSize_t + \alpha_3 SalesVolatility_{t-4,t} \\ & + \alpha_4 CashFlowVolatility_{t-4,t} + \alpha_5 OperCycle_{t-4,t} \\ & + \alpha_6 NationalAuditor_t + \alpha_7 \Delta SalesGrowth_t + \alpha_8 AbnormalReturn_t \\ & + \varepsilon_{t+1,t+4}. \end{aligned}$$

		Dependent Variable =			
	Pred.	***TotalEarnQuality***			
MgrlAbility	+	0.44***			0.38***
		6.07			*2.58*
Historical Ret	+		0.69***		0.83***
			12.29		*5.94*
Media Citations	+			−0.42***	−0.28*
				2.76	*−1.78*
FirmSize	+	0.08***	0.06***	0.08**	0.07*
		5.64	*3.72*	*2.23*	*1.92*
SalesVolatility	−	−0.49***	−0.73***	−0.66**	−0.87***
		−4.77	*−4.93*	*−2.19*	*−2.52*
CashFlowVolatility	−	−1.57***	−1.56***	−2.15***	−2.45***
		5.99	*−4.42*	*−2.55*	*−2.56*
OperCycle	−	−0.04	−0.07	−0.07	−0.07
		−1.29	*−1.86*	*−1.06*	*−1.01*
NationalAuditor	+	0.19***	0.04	0.33	0.06
		3.08	*0.56*	*1.13*	*0.22*
ΔSalesGrowth	?	−0.01*	−0.03*	0.01	0.04
		−1.65	*−1.75*	*0.21*	*0.65*
AbnormalReturn	?	0.09***	0.04*	0.17***	0.09**
		5.29	*1.82*	*4.53*	*2.19*
Total Observations		13,157	9,673	3,770	3,348
Pseudo R^2		0.80%	1.09%	0.40%	0.88%

*, **, *** Denote a two-tailed p-value of less than 0.10, 0.05, and 0.01, respectively.

This table reports the results from the ordered logistic regression of total earnings quality (*TotalEarnQuality*) on managerial ability and controls. *TotalEarnQuality* is the sum of three earnings quality variables: (1) the rank of estimation accruals quality ($ModifiedAQ_{Loss\%}$), (2) the rank of firm-specific earnings persistence, and (3) $-1 \times$ *Restate*. Thus, the variable ranges from a low of −1 to a high of 2. We decile rank *MgrlAbility*, *Historical Ret*, and *Media Citations* by industry-year. The sample size is reduced for specifications including *Media Citations* as this variable is limited to ExecuComp firms from 1995–2005. Intercept is included, but not tabulated. Z-statistics are presented in italics below the coefficients and are based on standard errors that are clustered by firm. We find similar results when we cluster by year.

See Table 1, Panel C for variable definitions.

matches specific managers with the years they managed the firm. The average manager fixed effect increases by 1.92 when moving from the lowest to the highest quartile of manager fixed effects. For comparison purposes, the average firm fixed effect increases by 2.25 when moving from the lowest to the highest quartile of firm fixed effects. Although fixed effects are quantifiable only for

CFOs switching firms within our sample, clearly manager-specific effects are economically significant.[33]

We next investigate if the CFO's score from his/her initial firm is associated with accruals quality at the subsequent firm. We expect firms that hire a more efficient CFO to experience an improvement in their earnings quality, and firms that hire a less efficient CFO to experience a decline in their earnings quality. Thus, using the sample of the 195 CFOs examined above who switched across our sample firms, we identify 116 with sufficient information to estimate the following:

$$\begin{aligned} \Delta ModifiedAQ_{Loss\%Ann} = {} & \alpha_0 + \boldsymbol{\alpha_1 \Delta MgrlAbility} + \alpha_2 \Delta FirmSize + \alpha_3 \Delta SalesVolatilty \\ & + \alpha_4 \Delta CashFlowVolatility + \alpha_5 \Delta OperCycle + \alpha_6 \Delta SalesGrowth_t \\ & + \alpha_7 \Delta AbnormalReturn_t + \varepsilon, \end{aligned} \tag{10}$$

where the change in earnings quality, as well as the change in each of the control variables, is measured from year $Variable_{c+1} - Variable_{c-1}$, where c is the year in which the CFO changed.[34] Thus, a positive value of $\Delta ModifiedAQ_{Loss\%Ann}$ signifies an improvement in earnings quality following the new CFO appointment.[35] The change in managerial ability reflects the difference between the newly appointed CFO's score from his or her prior firm and the departing CFO's score from the current firm, $MgrlAbility_{j,b,c-1} - MgrlAbility_{i,a,c-1}$, where manager b was hired by firm i and was previously employed by firm j. A positive value of $\Delta MgrlAbility$ signifies that the incoming manager is deemed more efficient than the outgoing manager. Thus, we expect α_1 to be positive. Table 10 reports the expected positive and significant coefficient on $\Delta MgrlAbility$. Finding that the association between ability and earnings quality spans firms helps to alleviate the general concern that the documented associations are attributable to the firm, rather than the manager.

Additional Control Variables

We consider two infrastructure-related control variables that have been shown to be associated with earnings quality: governance (Klein 2002) and internal control quality (Doyle et al. 2007). We measure governance with the percentage of independent board members, obtained from IRRC from 1996–2007, ranked by year and industry (*PctInd*), and proxy for internal control quality with the disclosure of material weaknesses in internal control (*ICW*). We obtain internal control data from Doyle et al. (2007) for 2002–2004 and from Audit Analytics from 2005–2007.[36] We find qualitatively similar results to those reported when models include these two additional controls.

We also explore the sensitivity of our results to using abnormal performance variables other than the change in sales growth. When we replace the change in sales growth with the level of sales growth and with historical three-year average change in sales growth, we find results similar to those reported.

VI. CONCLUSION

We examine the relation between managerial ability and earnings quality. While empirical literature in the area of earnings quality has largely focused on firm-specific characteristics, such as

33 Untabulated results indicate that the MA-Score is also positively correlated with the CFO-specific fixed effects.

34 We exclude *Loss%* as a control variable because it is used to identify $ModifiedAQ_{Loss\%}$. Results are similar if we include *Loss%* as an additional control variable.

35 As with the prior analysis, we use the annual accrual error, the absolute value of the error term from Equation (8), for this analysis to better match the accruals quality with the specific manager in place at that time.

36 The Doyle et al. (2007) data are available at: http://faculty.washington.edu/geweili/ICdata.html. We end the analysis in 2007 because of limitations in our board independence data.

TABLE 10

Change in Accruals Quality and Change in Managerial Ability

$$\Delta ModifiedAQ_{Loss\%Ann} = \alpha_0 + \boldsymbol{\alpha_1 \Delta MgrlAbility} + \alpha_2 \Delta FirmSize + \alpha_3 \Delta SalesVolatilty + \alpha_4 \Delta CashFlowVolatility + \alpha_5 \Delta OperCycle + \alpha_6 \Delta SalesGrowth + \alpha_7 \Delta AbnormalReturn + \varepsilon.$$

	Pred.	**Dependent Variable = $\Delta ModifiedAQ_{Loss\%Ann}$**
ΔMgrlAbility	+	0.04**
		1.98
ΔFirmSize	+	0.00
		0.49
ΔSalesVolatility	–	–0.02
		–0.77
ΔCashFlowVolatility	–	–0.02
		–0.19
ΔOperCycle	–	0.01
		0.39
ΔSalesGrowth	?	0.02***
		2.60
ΔAbnormalReturn	?	–0.01**
		–2.40
Total Observations		123
R^2		10.55%

, * Denote a two-tailed p-value of less than 0.05 and 0.01, respectively.
This table presents the OLS regression of changes in accruals quality on changes in managerial ability and changes in control variables. t-statistics are presented in italics below the coefficients. As our measure of managerial ability, we examine an annual version of the modified accruals quality measure examined in Table 8 (*Modified* $AQ_{Loss\%}$). The change in earnings quality, as well as the change in each of the control variables, is measured from year $Variable_{c+1} - Variable_{c-1}$, where c is the year in which the CFO changed. Thus, a positive value of $\Delta ModifiedAQ_{Loss\%Ann}$ signifies an improvement in earnings quality following the new CFO appointment. The change in managerial ability reflects the difference between the newly appointed CFO's score from his or her prior firm and the departing CFO's score from the current firm (i.e., $MgrlAbility_{j,b,c-1} - MgrlAbility_{i,a,c-1}$, where manager b was hired by firm i and was previously employed by firm j). A positive value of *ΔMgrlAbility* signifies that the incoming manager is deemed more efficient than the outgoing manager. We do not include *ΔNationalAuditor* as a control variable due to lack of variation. Intercept is included, but not tabulated.
See Table 1, Panel C for variable definitions.

firm size and board independence (Dechow and Dichev 2002; Klein 2002), we investigate manager-specific effects using the MA-Score, a measure of managerial ability developed in Demerjian et al. (2012). Our study is in the vein of Bertrand and Schoar (2003), who find that managers have an effect on firm choices such as acquisitions or research and development expenditures, and Francis et al. (2008), who find that earnings quality appears to vary with CEO reputation. Using the four alternative earnings quality measures of restatements, earnings persistence, error in the bad debt provision, and modified accruals quality, we find that more able managers report higher quality earnings.

Our study contributes to the earnings quality and managerial ability literatures. We document a positive association between managerial ability and earnings quality. We find that higher quality managers are associated with higher quality earnings. This finding is consistent with the premise that the more capable the manager, the better able he or she is to estimate accruals, and it suggests that firms can improve their earnings quality by employing higher ability managers. We also contribute to the earnings quality literature by offering a modification of the Dechow and Dichev accruals quality measure that more effectively partitions accruals quality from firms' underlying earnings processes.

Our study's first limitation is that our evidence relies on proxies for both managerial ability and earnings quality. Although we attempt to control for correlated variables such as economic shocks, the possibility of correlated omitted variables remains. Second, the ability score we examine is for the entire management team, while our focus in on the effect of CFOs and their delegates. We mitigate this concern to some extent by documenting economically significant CFO fixed effects on earnings quality—a methodology that does not rely on a specific managerial ability measure. We further corroborate our analysis by tracking CFOs across firms and documenting that hiring a higher ability CFO is associated with an improvement in the firm's earnings quality. Finally, our focus is on accrual estimation, and as such we abstract away from the decision to manage earnings. Future researchers might explore the association between managerial ability and earnings management, and similarly, the inter-relations between ability and infrastructure choices that have been shown to improve earnings quality, such as governance and internal control quality.

REFERENCES

Aier, J., J. Comprix, M. Gunlock, and D. Lee. 2005. The financial expertise of CFOs and accounting restatements. *Accounting Horizons* 19 (3): 123–135.

Anderson, K., and T. Yohn. 2002. *The Effect of 10-K Restatements on Firm Value, Information Asymmetries, and Investors' Reliance on Earnings*. Working paper, Georgetown University.

Ashbaugh-Skaife, H., D. Collins, W. Kinney, and R. LaFond. 2008. Internal control deficiencies, remediation and accrual quality. *The Accounting Review* 83 (1): 217–250.

Baik, B., D. Farber, and S. Lee. 2011. CEO ability and management earnings forecasts. *Contemporary Accounting Research* 28 (5): 1645–1668.

Ball, R., and L. Shivakumar. 2006. The role of accruals in asymmetrically timely gain and loss recognition. *Journal of Accounting Research* 44 (2): 207–242.

Bamber, L., J. Jiang, and I. Wang. 2010. What's my style? The influence of top managers on voluntary corporate financial disclosure. *The Accounting Review* 85 (4): 1131–1162.

Becker, C., M. DeFond, J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15 (1): 1–24.

Bertrand, M., and A. Schoar. 2003. Managing with style: The effect of managers on firm policies. *The Quarterly Journal of Economics*: 1169–1208.

Dechow, P. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics* 18: 3–42.

Dechow, P., and C. Schrand. 2004. *Earnings Quality*. Charlottesville, VA: The Research Foundation of CFA Institute.

Dechow, P., and D. Skinner. 2000. Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons* 14 (2): 235–250.

Dechow, P., and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (4): 35–59.

Dechow, P., W. Ge, and C. Schrand. 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics* 50: 344–401.

Dechow, P., and W. Ge. 2006. The persistence of earnings and cash flows and the role of special items: Implications for the accrual anomaly. *Review of Accounting Studies* 11 (2/3): 253–296.

DeJong, D., and Z. Ling. 2010. *Managers: Their Effects on Accruals and Firm Policies*. Working paper, The University of Iowa.

Demerjian, P., B. Lev, and S. McVay. 2012. Quantifying managerial ability: A new measure and validity tests. *Management Science* 58 (7): 1229–1248.

Doyle, J., W. Ge, and S. McVay. 2007. Accruals quality and internal control over financial reporting. *The Accounting Review* 82 (5): 1141–1170.

Dyreng, S., M. Hanlon, and E. Maydew. 2010. The effects of executives on corporate tax avoidance. *The Accounting Review* 85 (4): 1163–1189.

Fama, E., and K. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43 (2): 153–193.

Fee, C., and C. Hadlock. 2003. Raids, rewards, and reputations in the market for managerial talent. *Review of Financial Studies* 16 (4): 1315–1357.

Feroz, E., K. Park, and V. Pastena. 1991. The financial and market effects of the SEC's accounting and auditing enforcement releases. *Journal of Accounting Research* 29 (3): 107–142.

Francis, J., A. Huang, S. Rajgopal, and A. Zang. 2008. CEO reputation and earnings quality. *Contemporary Accounting Research* 25 (1): 109–147.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2004. Costs of equity and earnings attributes. *The Accounting Review* 79 (4): 967–1010.

Ge, W., D. Matsumoto, and J. Zhang. 2011. Do CFOs have style? An empirical investigation of the effect of individual CFOs on accounting practices. *Contemporary Accounting Research* 28 (4): 1141–1179.

Hennes, K., A. Leone, and B. Miller. 2008. The importance of distinguishing errors from irregularities in restatement research: The case of restatements and CEO/CFO turnover. *The Accounting Review* 83 (6): 1487–1519.

Hribar, P., and C. Nichols. 2007. The use of unsigned earnings quality measures in tests of earnings management. *Journal of Accounting Research* 45 (5): 1017–1053.

Jackson, S., and X. Liu. 2010. The allowance for uncollectible accounts, conservatism, and earnings management. *Journal of Accounting Research* 48 (3): 565–601.

Jones, K., G. Krishnan, and K. Melendrez. 2008. Do models of discretionary accruals detect actual cases of fraudulent and restated eamings? An empirical analysis. *Contemporary Accounting Research* 25 (2): 499–531.

Kim, Y., M. Park, and B. Wier, 2012. Is earnings quality associated with corporate social responsibility? *The Accounting Review* 87 (3): 761–796.

Klein, A. 2002. Audit committee, board of director characteristics and earnings management. *Journal of Accounting and Economics* 33 (3): 375–400.

Knechel, R., P. Rouse, and C. Schelleman. 2009. A modified audit production framework: Evaluating the relative efficiency of audit engagements. *The Accounting Review* 84 (5): 1607–1638.

LaFond, R. 2008. Discussion of: CEO reputation and earnings quality. *Contemporary Accounting Research* 25 (1): 149–156.

Lev, B., and T. Sougiannis. 1996. The capitalization, amortization, and value-relevance of R&D. *Journal of Accounting and Economics* 21 (1): 107–138.

Libby, R., and J. Luft. 1993. Determinants of judgment performance in accounting settings: Ability, knowledge, motivation and environment. *Accounting, Organizations and Society* 18 (5): 425–450.

McNichols, M. 2002. Discussion of: The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (4): 61–69.

McNichols, M., and P. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (Supplement): 1–31.

Milbourn, T. 2003. CEO reputation and stock-based compensation. *Journal of Financial Economics* 68: 233–262.

Palmrose, Z.-V., and S. Shultz. 2004. The circumstances and legal consequences of non-GAAP reporting: Evidence from restatements. *Contemporary Accounting Research* 21 (1): 139–180.

Palmrose, Z.-V., V. Richardson, and S. Shultz. 2004. Determinants of market reactions to restatement announcements. *Journal of Accounting and Economics* 37 (1): 59–89.
Petersen, M. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22 (1): 435–480.
Plumlee, M., and T. Yohn. 2010. An analysis of the underlying causes attributed to restatements. *Accounting Horizons* 24 (1): 41–64.
Richardson, S., R. Sloan, M. Soliman, and I. Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39 (3): 437–485.
Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289–315.
Tan, H., and K. Jamal. 2006. Effect of accounting discretion on ability of managers to smooth earnings. *Journal of Accounting and Public Policy* 25: 554–573.
Wysocki, P. 2009. *Assessing Earnings and Accruals Quality: U.S. and International Evidence*. Working paper, University of Miami.

APPENDIX A
ESTIMATION OF TOTAL FIRM EFFICIENCY
(Demerjian et al. 2012)

Our main measure of managerial ability, the MA-Score, is the metric developed by Demerjian et al. (2012), using data envelopment analysis (DEA), a nonlinear optimization program that calculates unit-specific relative efficiency. The program is as follows:

$$\max_{v,u}\theta = \frac{\sum_{i=1}^{s} u_i y_{ik}}{\sum_{j=1}^{m} v_j x_{jk}}. \tag{A1}$$

Subject to:

$$\frac{\sum_{i=1}^{s} u_i y_{ik}}{\sum_{j=1}^{m} v_j x_{jk}} \le 1 \;\; (k = 1, \ldots, n); \tag{A2}$$

$$v_1, v_2, \ldots, v_m \ge 0; \tag{A3}$$

$$u_1, u_2, \ldots, u_s \ge 0. \tag{A4}$$

DEA measures the efficiency of a single unit, here firm k, relative to a set of comparable firms. The objective function measures efficiency as the weighted outputs scaled by the weighted inputs. There are s outputs and m inputs, indexed by i and j, respectively. The quantities of each output i and input j for firm k are y_{ik} and x_{jk}, respectively. The optimization program maximizes (A1) by selecting the weights on each output (u_i) and input (v_j). The vectors of weights on the outputs (u) and inputs (v) are termed implicit weights. Efficiency is based on the level of the weighted outputs to the level of the weighted inputs. The most efficient firms have the highest level of outputs for a fixed level of inputs (or equivalently, the lowest level of inputs for a fixed level of outputs). DEA calculates a unique set of implicit weights for each firm k.

The first constraint, (A2), scales the implicit weights so that the most efficient firm (or firms) has (have) an efficiency value of 1. The optimal weights for each firm k are tested for all the other

comparable firms (1,. . .,*n*; $\neq k$). This calculates what the efficiency would be for each comparable firm under the implicit weights calculated in (A1) for firm *k*, allowing for the determination of relative efficiency. Constraints (A3) and (A4) require implicit weights to be non-negative, which prevents solutions calling for negative input levels.

Total firm efficiency is estimated using a single output and seven inputs. Total revenue ("SALE") is the output, as the principal objective of the firm is to produce sales. The most successful firms are those that produce the maximum sales at the lowest cost. The cost of producing the sales is captured by the seven inputs. The first five correspond to assets the company invests in that are expected to affect their revenue-generation. Demerjian et al. (2012) consider the beginning-of-period balance for each of these assets, since managers' past decisions regarding these assets are expected to affect current-period revenues.

1. Net Property, Plant, and Equipment (PP&E; "PPENT").
2. Capitalized Operating Leases. The discounted present value of the next five years of required operating lease payments (available in the firm's footnotes to the financial statements and on Compustat).[37] The inclusion of Net Operating Leases as an input increases the input comparability among firms that effectively have the same operations, but either lease or buy their revenue-generating equipment.
3. Net Research and Development (R&D). To calculate net R&D, which is not reported as an asset on the balance sheet, Demerjian et al. (2012) follow Lev and Sougiannis (1996), who use a five-year capitalization period of R&D expense ("XRD"), where the net value (net of amortization) is:

$$RD_{cap} = \sum_{t=-4}^{0} (1 + 0.2t) \times RD_{exp}.$$

   Thus, for example, R&D expenditures from five years back receive a weight of 0.2 (they were already amortized 80 percent), four years back receive a weight of 0.4 (amortized 60 percent), etc., with the prior year's R&D ($t = -1$) receiving full weight.
4. Purchased goodwill, reported on the balance sheet, which is the premium paid over the fair value of a business acquisition ("GDWL"). Goodwill generally reflects the value of the acquired intangible assets.
5. Other acquired and capitalized intangibles ("INTAN" less "GDWL"), also reported on the balance sheet, which includes items such as client lists, patent costs, and copyrights.

They also include two year *t* expenses, Cost of Goods Sold and Selling General and Administrative Expense, to account for the cost of inventory (Cost of Goods Sold) and sales generated from advertising and the quality of the sales force (advertising, training costs, and IT services are included in SG&A).

They estimate DEA efficiency by industry, based on Fama and French (1997), to increase the likelihood that the peer firms have similar business models and cost structures within the estimations. The resulting score ranges from 0–1, with 1 being the optimal output for a given mix of inputs.

Using DEA instead of traditional ratio analysis has several advantages. First, DEA allows the weightings on each of the inputs to vary, whereas traditional efficiency ratios restrict all weightings to be equal to 1. For example, within DEA, one dollar at historical cost, such as PP&E, can count

[37] Demerjian et al. (2012) use a discount rate of 10 percent per year. Data items for the five lease obligations are "MRC1–MRC5." They note that they would also like to discount the "thereafter" payments; however, this line item was not collected by Compustat for the bulk of the sample period. Note that capital leases are included in Net PP&E.

differently from one dollar at or near replacement cost, such as COGS, but both are weighted identically in a traditional efficiency ratio. Second, DEA compares each firm within an industry to the most efficient firm, whereas traditional efficiency analysis compares each firm to the mean or median firm. Demerjian et al. (2012) provide additional information, explicit comparisons of this score with a residual from an OLS regression, and a comparison of variable ratios, such as return on assets.

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 499–520

American Accounting Association
DOI: 10.2308/accr-50319

The Effect of the Timing and Direction of Capital Gain Tax Changes on Investment in Risky Assets

Diana Falsetta
University of Miami

Timothy J. Rupert
Arnold M. Wright
Northeastern University

ABSTRACT: This study examines the effect of timing (gradual versus immediate) and direction (tax increase or decrease) of a tax change on taxpayer behavior. Specifically, we focus on capital gain tax changes and preferences for investment in riskier assets. We run an experiment with 117 participants who allocate investment dollars between two funds of differing risk. Drawing on mental accounting and hedonic editing (Thaler 1985; Thaler and Johnson 1990), we posit that a tax decrease (a "gain") implemented gradually over several years will result in a greater increase in risky investment once the decrease is fully implemented than when the tax change is implemented all at once. In contrast, once a tax increase (a "loss") is fully implemented, a smaller decrease in risky investment will result when the change occurs all at once rather than gradually. Our findings support these expectations, suggesting that the manner of implementing a tax law change may impact decisions.

Keywords: *capital gains; hedonic editing; investments; mental accounting; prospect theory; tax.*

Data Availability: *Contact the authors.*

We thank Andy Cuccia, Amy Hageman, Brian Hogan, Janet Meade, Michaele Morrow, Jennifer Schafer, Cynthia Vines, and Amber Whisenhunt, as well as participants at the 2010 *JATA* Conference, the 2010 American Accounting Association (AAA) Annual Meeting, the 2010 Accounting, Behavior and Organizations Research Conference, the 2010 South Florida Accounting Research Conference, and the research seminars at Florida International University and Oregon State University for their helpful comments and suggestions on earlier versions of the manuscript.

Editor's note: Accepted by John Harry Evans III, with thanks to Steven Kachelmeier for serving as editor on a previous version.

Submitted: August 2010
Accepted: September 2012
Published Online: October 2012

I. INTRODUCTION

Congress frequently changes the Internal Revenue Code. In her 2008 annual report to Congress, the National Taxpayer Advocate, Nina Olson, noted that there were more than 3,250 changes in the Internal Revenue Code between 2001 and 2008 (National Taxpayer Advocate 2008, 4). In enacting these tax changes, Congress has often incorporated transition rules in the form of phase-ins, delayed effective dates, or grandfathering (Logue 1996). The primary stated reason for transitioning the law is to allow taxpayers time to re-structure their transactions in order to minimize unexpected consequences from the tax change. In addition, these transition rules have sometimes been used to lessen the budget impact of the tax change.

However, the timing of changes in tax policy has not been without controversy. Some policy experts have argued that timing policies are necessary because they protect taxpayers who have relied on existing laws for decisions they have made (Goldberg 1994; Logue 1996). In contrast, others have argued that timing rules are not necessary because taxpayers can anticipate when changes will occur (Graetz 1977; Kaplow 1986). In practice, Congress has used a variety of timing methods to transition from one tax structure to another. In some cases, Congress has fully implemented changes immediately with the enactment of the law, while in other situations legislators have implemented the change gradually by using phase-ins, delayed effective dates, or grandfathering.[1] While these gradual changes may add to the complexity of the law, Congress has often stated that they are justified in helping taxpayers cope with the change.

Given the variety of choices for the timing of tax changes, an understanding of how taxpayers react to differing implementation strategies is important. However, to date, little research has been conducted to examine how taxpayers respond to variations in the timing of tax changes. The present study investigates the impact of both the timing and the direction of a tax change to better understand these effects.

Using an experiment, we examine whether taxpayers change their preferences for risky investments in response to both the timing (gradual versus immediate) and the direction (tax increase or decrease) of a capital gain tax change. Relying on mental accounting and the hedonic editing hypothesis (Thaler 1985; Thaler and Johnson 1990) as a theoretical framework, we posit that taxpayers will invest more in a riskier investment when there is a capital gain tax decrease that is implemented gradually rather than all at once. This difference is posited because a tax decrease is viewed as a "gain" to taxpayers, and the theory predicts that it is preferable to experience several smaller gains than one large gain. We hypothesize the opposite effect for a capital gain tax increase (a "loss") and expect a greater reduction in riskier investment for an immediate change rather than a gradual change.

We run an experiment with 117 participants who have experience with investment and tax-reporting decisions. We randomly assigned the participants to four treatments, as well as to a control group with no tax effects. In the four treatment groups, participants experienced a tax change that was either gradual or immediate and involved either a tax rate increase or decrease. Following the decision task used by Meade (1990), participants allocated investment dollars between two funds of differing risk over seven periods. The findings support our hypotheses and suggest that the timing of a tax change has important implications for taxpayer decisions.

---

[1] Examples of immediate tax decreases include the reductions in the tax rate on capital gains and dividends that Congress passed in 2001 and 2003. Congress has gradually phased-in decreases in tax with the elimination of the estate tax and the elimination of the phase-outs of personal exemptions and itemized deductions for higher-income taxpayers. While tax increases have been less common over the last decade, Congress immediately implemented a tax increase by raising the age at which taxpayers are subject to the "kiddie" tax in 2008. Gradual tax increases through phase-outs were very common with the elimination of deductions through the Tax Reform Act of 1986 (e.g., the elimination of the deduction for personal interest).

Section II next reviews the literature related to decision making and taxes, and poses our hypotheses related to the timing and direction of the tax change. Section III of the paper describes the experimental methods used to examine the hypotheses. The results follow in Section IV. Section V offers conclusions and implications of the findings for tax policy and future research.

II. LITERATURE REVIEW AND HYPOTHESES

Academicians and politicians have long been interested in the effects of tax policy on the willingness of taxpayers to invest in risky assets. Early studies used models of taxpayer behavior to examine the effects of proportional taxes (Stiglitz 1969; Domar and Musgrave 1944) and progressive taxes (Schneider 1980; Fellingham and Wolfson 1978) on risky investment. More recent studies have used experimental methods to address this issue (Anderson and Butler 1997; Meade 1990; King and Wallin 1990).

Politicians frequently debate the preferred tax policy to encourage investment in risky assets (Lerner 2002; Dunn 1998).[2] The debates are reflected in the frequent changes in various provisions related to these investments. In particular, Congress has frequently changed the capital gain and loss provisions over the last two decades. Recent tax law changes suggest that these provisions will continue to be the focus of much attention, making it an appropriate setting for an investigation of the effects of tax law changes.

The effect of changes in the taxation on the sale of capital assets depends on a variety of attributes of the tax system, including the tax rate that applies to capital gains and the deductibility of capital losses. When tax rates are reduced on capital gains and capital losses are fully deductible, investment in risky assets are expected to increase (Burman 1999). Because a decrease in tax rates increases the difference in expected after-tax returns between riskier and less risky assets, as we illustrate next, the riskier asset will become more attractive to investors.[3] As in other studies using the standard economic approach, we assume taxpayers are generally risk-averse.

Consider an example where the taxpayer has a choice between two investments—a riskier investment with a pre-tax expected return of 60 percent, and a less risky investment with a pre-tax expected return of 25 percent. Assuming the earnings from both investments would be taxed as a capital gain at a rate of 35 percent, the taxpayer would be comparing an expected after-tax return of 39.0 percent ($60\% \times [1 - 0.35]$) for the riskier investment versus 16.25 percent ($25\% \times [1 - 0.35]$) for the less risky investment, resulting in a 22.75 percentage point difference in after-tax expected return.

If Congress reduces the tax rate on capital gains, then the difference in expected after-tax returns for the two investments will increase. Assuming that Congress lowered the tax rate on capital gains from 35 percent to 20 percent, the after-tax return on the riskier investment would increase to 48.0 percent ($60\% \times [1 - 0.20]$), while the after-tax return on the less risky investment would increase to 20.0% ($25\% \times [1 - 0.20]$). Thus, the difference between the after-tax expected returns from the two investment options would increase to 28.0 percentage points. Because the pre-tax expected returns of the two investment options has not changed with the decrease in tax rate, the riskier investment will be more attractive to investors than it was previously.

---

[2] The motive(s) of Congress may vary for tax changes. For instance, provisions may change because of differences in views regarding *how* to best encourage investment in risky assets or *whether* to do so. Our focus in the present study is on the former as this seems to be more pervasive in recent years.

[3] We recognize, however, that an increase in the capital gains rate may, in some circumstances, positively affect an investor's expected after-tax returns as a result of lowering risk (Sikes and Verrecchia 2012). However, in our experimental case none of these conditions are present (i.e., a very high investment systematic risk, a very high market risk premium, or a very low risk-free rate of return).

In examining the effect of taxes on risky investment, much of the previous literature has been built on the standard economic approach demonstrated in this example. When making a decision under risk, taxpayers are generally price-takers who cannot alter returns by making a particular investment, and are expected to choose the alternative that affords the greatest expected utility (von Neumann and Morgenstern 1944). This expectation assumes that individuals conform to the tenets of economic "rationality." However, as with many decisions involving choice under uncertainty, individuals frequently depart from maximizing expected utility and display biases and limits on their ability to compute expected utility. For example, many of the early studies on taxpayer evasion used the expected utility model to determine how taxpayers respond to increased audit and penalty rates and what is the optimal enforcement strategy (e.g., Allingham and Sandmo 1972; see Cowell [1990] for a review). However, researchers have emphasized that, given the existing low audit rate, expected utility theory implies that rational individuals should pay far less in taxes than they actually do. The limits of using expected utility theory to explain individual decisions have led to the examination of other theories. In particular, we next discuss mental accounting and the hedonic editing hypothesis.

Prospect Theory

Mental accounting and the hedonic editing hypothesis have their foundations in prospect theory's S-shaped value function (Kahneman and Tversky 1979). Thaler (1999, 185) states, "The value function can be thought of as a representation of some central components of the human perceived pleasure machine. It has three important features each of which captures an essential element of mental accounting." First, the value function is defined in terms of gains or losses, relative to some neutral reference point—usually the *status quo*—rather than in absolute terms. Second, the value function is concave for gains and convex for losses. This represents the concept of diminishing sensitivity, where each additional unit of gain adds less value or pleasure than the preceding one, and each unit of loss causes a smaller decrease in value or pain than the preceding one. Individuals are more sensitive to gains or losses near the *status quo* than to equal gains or losses farther from the *status quo*. Third, the value function is steeper for losses than for gains, so that losses yield more pain than the pleasure of an equivalent gain.

Mental Accounting and the Hedonic Editing Hypothesis

Building on prospect theory, which considers decision making under uncertainty with respect to *single* outcomes, mental accounting (Thaler 1985) provides an alternative set of expectations by considering the effect of *multiple* outcomes (i.e., x and y). Mental accounting predicts that the manner in which people respond to outcomes involving multiple events depends on whether the events are represented as belonging together in a single mental account or separately in different mental accounts (Kahneman and Tversky 1984; and Tversky and Kahneman 1981). Seeking to understand how individuals combine or code *multiple* events, using prospect theory's value function, Thaler and Johnson (1990) propose that individuals edit their choices according to various rules, often in a manner that would make their choices most pleasant or least unpleasant. These rules, which are referred to as the hedonic editing hypothesis, follow four principles that state that for a joint outcome (x, y) individuals' value functions, $v(\cdot)$, exhibit a preference to: (1) segregate gains or pleasant outcomes $[v(x) + v(y) > v(x + y)]$, (2) integrate losses or unpleasant outcomes $[v(-(x + y)) > v(-x) + (-y)]$, (3) integrate smaller losses with larger gains, and (4) segregate small gains from larger losses, where $v(x)$ and $v(y)$ are gains and $v(-x)$ and $v(-y)$ are losses.

Mental accounting and hedonic editing characterize decision makers as value maximizers who mentally segregate or integrate outcomes depending on which mental representation is more desirable. Individuals are expected to segregate gains and integrate losses because the value

function exhibits diminishing sensitivity as the magnitude of a gain or a loss becomes greater. In particular, individuals can maximize their happiness by savoring gains one by one and minimize the pain by facing one large loss rather than equivalent, smaller individual losses. Let us assume an individual is facing potential multiple losses or gains. Because the value function is convex (concave) in the domain of losses (gains), value is maximized by combining (segregating) two losses (gains) rather than segregating (combining) them, since the second loss (gain) is experienced on the flatter, or less sensitive, part of the curve.

The core of the hedonic editing rules focuses on the segregation and integration of outcomes so as to maximize value. However, they also provide support for the way in which the segregation or integration impacts risk-taking. In particular, Thaler and Johnson (1990) examined how risk-taking behavior is affected by prior gains and losses, which involve preferences for whether events would occur on the same day or different days. They observe increased risk-seeking (and "more happiness") in the presence of a prior gain ("house money effect"). In addition, they observe risk-aversion in the presence of a prior loss when not offered the opportunity to break even.

Thaler and Johnson (1990) propose that temporal separation of events facilitates segregation and temporal proximity facilitates integration.[4] The hedonic editing rules imply that people prefer to experience events on different days when segregation is preferred (i.e., gains) and on the same day when integration is preferred (i.e., losses). Thaler and Johnson (1990) and Linville and Fischer (1991) provide support for the tendency to segregate gains, although the integration of losses is not supported by these studies. In more recent research, Lim (2006) and Arkes et al. (2008) find support for both the segregation of gains and integration of losses. Specifically, Lim (2006) examines the trading records of individual investors at a large discount brokerage firm and finds that investors are more likely to bundle the sale of their loss stocks on the same trading day than the sale of their gain stocks.

According to the hedonic editing rules, individuals prefer to spread out the arrival of pleasant events (e.g., a tax rate decrease), to help accentuate the pleasures experienced. In contrast, individuals prefer to combine the arrival of unpleasant events (e.g., a tax rate increase), to reduce the unpleasant experiences.

Ceteris paribus, following the traditional arguments about the effects of changes in the tax rates on capital gains, a decrease in the taxpayer's tax rate is expected to result in a greater increase in risky investments, because of the resulting higher expected after-tax returns, which stimulate investment. Likewise, an increase in the taxpayer's tax rate is expected to result in a decrease in risky investments, because of the resulting lower expected after-tax returns, which inhibit investment. Previous research on the level of capital gain taxation has generally supported this assertion (Meade 1990). However, previous research has neither focused on the process of the tax change nor its consequences. When these factors are considered, mental accounting theories predict different outcomes.

The Timing of the Tax Change

Congress has a variety of timing methods for implementing tax changes (e.g., legislating a 21 percent decrease in the capital gains tax rate could be implemented all in one year or 7 percent for each of the successive three years). Given the capital gain tax rate incentives, under standard economic theory, taxpayers facing a tax rate decrease should increase their investment in risky assets, while taxpayers facing a tax rate increase should decrease their investment in risky assets, regardless of the method of implementation. However, mental accounting theory and the hedonic

---

[4] While prior research has used the editing rules to explain prospective evaluations only, recent literature has shown that it may also be used for retrospective evaluations of experiences (Cowley 2008).

editing hypothesis predict that taxpayers will prefer the method of timing that makes them the happiest. As noted in the previous section, this literature suggests that segregating gains and integrating losses will result in greater happiness. As a result, those who are happy (unhappy) will increase (decrease) their risk-taking. This response is based on the role of affect (pleasant/unpleasant feelings) on participants' willingness to take risk.

Affect includes the emotional response, including feelings or mood, to a positive/negative event. Various studies have examined the role of affect and decision frames (gains or losses) and risk-taking (e.g., Seo et al. 2010; Au et al. 2003; Isen and Patrick 1983). Specifically, this research has found that positive affect promotes risk-seeking behavior. These studies have induced positive (negative) affect in various ways, such as providing for an experimental gain/loss condition that results from a stock price increase (decrease), similar to our tax rate decrease (increase) conditions. Participants' affective states (feelings of pleasantness) are then measured in response to the gain/loss. After experiencing a gain (loss), participants reported pleasant (unpleasant) feelings and subsequently reported a greater propensity to increase (decrease) risk-taking.

Using an Internet trading experiment, Au et al. (2003) found that traders experiencing pleasant feelings placed larger bets, while those with unpleasant feelings made more conservative choices. In addition, Seo et al. (2010) note that affect influences risk-taking via its effect on cognitive processes, and that pleasant/unpleasant feelings interact with situational context in a complex manner. They found that pleasant feelings led to increased risk-taking in the realm of gains, but decreased risk-taking in the realm of losses.

Considering a tax rate decrease (a gain for the taxpayer), we posit that because of the concavity of the prospect theory value function for gains, taxpayers are likely to segregate the tax rate cuts if Congress implements the tax decrease gradually over a period of time. Because of the steepness of the value function near the reference point and its flatness further away from the reference point, taxpayers are expected to experience greater sensitivity to, and pleasant feelings with, three smaller tax cuts than one larger tax cut. Because of the greater value and pleasant feelings experienced with the three smaller gains (i.e., the smaller gains seem larger to the taxpayer than one large gain even though they are equivalent), we posit that based on mental accounting and hedonic editing, the increase in the riskier investment that should be associated with a capital gains tax rate cut will be larger when the gain is experienced gradually in smaller decreases rather than immediately as one large decrease.

Similarly, when faced with a tax rate increase (a loss for the taxpayer), based on mental accounting and hedonic editing, we posit that taxpayers are likely to prefer to integrate these losses rather than experience them as segregated losses. Due to the steepness of the value function near the reference point, we predict that three successive smaller tax increases will cause taxpayers to experience a greater feeling of loss and unpleasantness than will one large tax increase. Thus, we posit that the decrease in riskier investment that is expected to be associated with a capital gains tax rate increase will be greater with the gradually implemented tax increases rather than with the immediately implemented tax increase.[5]

Thus, even though the tax benefit/cost is the same under either implementation method once the tax change is fully implemented, taxpayers will prefer the implementation method that mentally produces the most pleasant outcome. Therefore, we posit an ordinal interaction, as reflected in the following hypothesis:

[5] While in practice the time value of money may be a consideration by taxpayers, this is unlikely to have an impact in an experiment where the duration of decisions is over a short time period. Consideration of this factor would serve to bias against supporting our hypothesis where we predict taxpayers prefer a gradual (immediate) tax decrease (increase), which would lower the present value of such returns versus an immediate (gradual) tax decrease (increase).

H1: The interaction of the timing and the direction of the tax change will have a significant impact on the change in amount invested in the risky investment once the change is fully implemented.

More specifically, we predict the following two subsidiary hypotheses for the tax increase and tax decrease situations:

H1A: Because individuals prefer to segregate gains, a tax *decrease* that is implemented gradually over several periods will result in a greater increase in risky investment once the decrease is fully implemented than when the tax change is implemented all at once.

H1B: Because individuals prefer to integrate losses, a tax *increase* that is implemented gradually over several periods will result in a greater decrease in risky investment once the increase is fully implemented than when the tax change is implemented all at once.

III. METHOD

We test the research hypotheses using an experiment, administered through a website, with a 2 $\times$ 2 between-subjects design involving the "Timing" and "Direction" of the tax change.[6] Timing is manipulated at two levels with either a gradual phase-in over three periods or an immediate phase-in during a single period, while the Direction of the change is either a tax increase or a tax decrease. In terms of the predictions of the hedonic editing rules, a tax increase is viewed as a "loss" and a tax decrease as a "gain." The dependent variable is the change in the allocation of available funds to the riskier asset after the implementation of the tax change.

In addition to the four treatment groups, we also include a fifth control group with the same decision parameters as the treatment groups except that no tax changes are present. This group of participants allows us to control for the potential effects of prior period outcomes on investment decisions (e.g., a high return for the riskier investment in the previous year) and enables us to normalize responses relative to the control group.

Experimental Task

To investigate the research hypotheses, we use a task patterned after the investment choice designed by Meade (1990). Participants allocate an available investment pool between two mutual funds of different risk: Fund A, which is lower risk, and Fund B, which is higher risk. We provide a choice between two funds, since investment decisions always inherently involve two or more alternatives such as investing in a mutual fund, purchasing common stock, or depositing money in a savings account. Exhibit 1 shows the expected returns and probabilities for the two mutual funds.[7]

To simplify the task setting and control potential confounding factors, we hold a number of variables constant in the experiment. Mutual fund returns and probabilities are the same across

---

[6] Much of the prior literature in mental accounting has had participants make a choice between two alternatives (within-subjects), whereas our study elicited preferences separately (between-subjects). We chose this design because it is consistent with the existing tax system. In general, taxpayers are not given a choice on whether a tax change is implemented gradually or immediately. Congress decides how to implement a tax change and imposes it on taxpayers. This feature is consistent with a between-subjects design, i.e., examining responses given a particular context. A recent study examining within- versus between-subjects designs in mental accounting experiments (Chatterjee et al. 2009) found that within-subjects designs can often exacerbate or strengthen biases, instead of reducing judgmental biases.

[7] We use the same rates of return as in Meade (1990), who based the rates of return in her study on historical returns for the Dow Jones indices for the less risky fund and the three largest publicly held venture capital investment companies for the more risky fund. While these rates may not be reflective of current returns, we held this information constant across all treatment conditions.

EXHIBIT 1
Parameters of the Investment Decision and Tax Rate Schedule Used in the Experiment

Investment Decision:

Mutual Fund A (lower risk investment fund):
Probability of Investment Gain/Loss:
15% chance of a 0% gain
40% chance of a 20% gain
25% chance of a 35% gain
20% chance of a 60% gain

Mutual Fund B (higher risk investment):
Probability of Investment Gain/Loss:
1% chance of a 100% loss
14% chance of a 25% loss
50% chance of a 50% gain
20% chance of a 100% gain
15% chance of a 150% gain

Actual Outcomes for each Period:

	Mutual Fund A	Mutual Fund B
Period 1	60%	50%
Period 2	35%	50%
Period 3	60%	100%
Period 4	60%	50%
Period 5	20%	50%
Period 6	35%	100%
Period 7	20%	150%

Tax Rate Schedule:

Taxable Income Over	But Not Over	Base Tax + Tax Rate
$ 0	$ 20,000	$ 0 + 15.0%
20,000	60,000	3,000 + 27.0%
60,000	150,000	13,800 + 36.0%
150,000	—	46,200 + 40.0%

periods. The information about returns and probabilities is disclosed to participants during each period in the form of two pie graphs that present the returns and probabilities. The funds available for investment are also held constant across experimental conditions and across experimental periods, with participants informed that they have $20,000 to invest each period. Finally, participants are given a salary amount each period ($101,750) that does not change during the experiment.

Individuals may link anticipated returns and, hence, decisions to invest in various options to the assumed state of economic trends. For instance, in a growth economy one may expect strong returns across many industries, fostering a choice to invest in riskier options. Further, certain sectors, such as retail and travel, traditionally benefit more than other sectors from a strong economy. To control for this potentially confounding effect, participants are told that projections indicate a relatively stable economy with average growth of 2 to 3 percent over the period of the study.

The actual outcomes for each fund each period are also included in Exhibit 1 and were the same for all participants.[8] These outcomes are independent of the individual's investment decisions and the tax system. After participants allocate their investment dollars to the two funds in each period, the program then provides the investment outcomes for that period and calculates a summary of the investment results based on the allocation of funds, providing participants with an opportunity in the next period to consider adjustments in the allocation among the investment choices. This summary includes the participant's investment earnings, taxes on earnings, and after-tax earnings.

To avoid "end-of-game" effects, participants are not told the number of investment periods at the start of the experiment. However, the experiment lasted for seven periods. Periods 1 and 2 are "base" cases with no tax change and serve as a within-subject control group. Participants then make investment choices for the next three periods in which, for the treatment groups, there is either a tax increase or decrease that occurs either all at the beginning of Period 3 or is phased-in gradually over Periods 3, 4, and 5.

In addition to the information regarding the probability of a gain or loss on the different investment options, participants are also provided with information about the tax system. Participants are informed that the experimental tax rate structure is designed to be similar to the current progressive federal tax rate structure. Exhibit 1 shows that the tax rate structure ranges from a marginal tax rate of 15 percent to 40 percent. To make the tax structure salient and simplify the task, the salary income for each period is chosen so that participants fall into the second-to-highest tax bracket throughout the experiment, so the marginal tax rate on ordinary income remains constant at 36 percent.

We informed participants that investment gains are taxed at a separate rate.[9] All treatment group participants are then told: "At different points in time, Congress has varied the investment gains tax rate, making it either lower than or the same as the ordinary income tax rate. Generally, Congress lowers the rate on capital gains to encourage greater investment in risky assets necessary for economic growth, while it increases the tax on capital gains to raise revenue." The purpose of indicating Congressional motives was to provide a context for participants; as noted, the same information was provided to all participants.[10] We then told participants in the tax decrease condition that the tax rate on investment gains is currently 36 percent. In the increase conditions, we informed participants that the current tax rate on investment gains is 15 percent. In the introduction to the experimental periods, we provided participants with an example of how both ordinary income and investment gains are taxed to ensure that they are familiar with the tax system. Participants are also reminded of the tax rate on both ordinary income and investment gains each period.

---

[8] We determined the actual outcomes for each period by performing random draws from the distribution of likely outcomes given to the participants. All participants experienced the same outcome for any particular period.

[9] Unlike the current federal tax system, there was no limitation on the deductibility of capital losses in the experiment. Limitations on the deductibility of net capital losses may affect taxpayers' willingness to invest in risky assets. However, participants in both the tax increase and tax decrease conditions faced the same regime on loss deductibility, so while the lack of a limitation on the deductibility of losses may influence the overall level of investment in risky assets, it should not affect our comparison of the change in riskier investment between the gradual and immediate conditions.

[10] At the individual taxpayer level, it may not always be desirable to increase investment in riskier assets (for example, it may not be optimal for retirees to substantially invest in such assets). However, in advocating cuts in the capital gains tax rate, the impact on investment in riskier assets is often a key consideration by Congress. For example, when arguing for a decrease in the capital gains rate in 1997, Senator Orrin Hatch noted that "the capital gains tax is a tax on success. It is an additional tax on the reward for taking risks. The American dream is not dead; it's just that we have been taxing it away" (*Tax Notes Today* 1997). In our experimental task, participants are given the scenario that they are investing available cash. So while Congress may not want all taxpayers to increase investment in riskier assets, taxpayers in this situation are likely among those whom Congress is hoping will increase investment in riskier assets in an effort to foster economic growth.

After two periods with the tax system in place, the participants received information about a tax change. In the two tax increase treatments, we informed the participants that to raise revenue Congress decided to increase the tax rate on investment gains or losses from 15 to 36 percent. In the immediate tax increase condition, we told the participants that the 21 percent increase will take effect immediately. In the gradual increase condition, we told the participants that Congress has decided to increase the rate gradually over the next three periods by 7 percent each period.[11]

In the two tax rate decrease treatments we told participants that to encourage greater investment in risky assets Congress has decided to decrease the tax rate on investment gains or losses from 36 to 15 percent. In the immediate tax decrease condition, we told participants that the 21 percent decrease will take effect immediately. In the gradual decrease condition, we told them that Congress has decided to implement the decrease gradually by reducing the rate by 7 percent in each of the next three periods.

One potential concern is that participants may have differing expectations regarding the stability in the tax system. To deal with this issue at the onset of the experiment, we told participants that the changes are guaranteed by Congress for a three-year period after which they will be re-evaluated by Congress.[12]

After the three-year period during which the tax rate change is guaranteed, the program presents the participants with a message that Congress has decided to extend the rate change for an additional three periods. After receiving this message, participants complete two additional periods and then answer manipulation-check questions, along with background and demographic questions. Exhibit 2 provides a summary of the tax rate on investment gains that participants experienced in each of the experimental periods.

In summary, investors make a new investment decision each period and are aware of the nature of the tax change in Period 3 and how it will be implemented. Our theoretical expectations are that a tax decrease (increase) is viewed by an investor as a gain (loss). Thus, the mental accounting begins in Period 3 where a gradual phase-in of a gain (loss) is viewed as pleasant (unpleasant) versus an immediate phase-in. As such, we predict that the gradual phase-in will lead to more (less) investment in the riskier asset after the tax change is fully implemented at the end of Period 5 for the gain (loss) than under the immediate phase-in.

Participants and Incentives

Participants for the study are M.B.A. students recruited at two large universities. Of 264 total students enrolled in the classes selected for recruitment, 141 responded by logging onto the website for the study, but 15 of the 141 did not complete the experimental task. We dropped two participants because of a technical problem with the recording of their responses and seven participants because they responded incorrectly to two manipulation-check questions. Thus, the final sample size was 117 for a response rate of 44.3 percent.

Because the experimental task involves consideration of tax effects in making investment choices, it is important that participants have familiarity with tax consequences as well as investment decisions. Table 1 shows that participants, on average, filed federal tax returns 3.68 times in the past five years and almost half of the respondents (46.2 percent) prepared their own

[11] While identification of the purpose of the tax change may potentially impact participants' level of investment in the riskier asset, this information is held constant for the two tax increase and two tax decrease conditions. Our hypotheses focus on the levels of investment within each timing condition (gradual versus immediate).

[12] An alternative is to put the experiment in the context of a "management fee" rather than the tax system (Meade 1990). The disadvantage of this alternative is that it then removes the participant from the normal tax context, raising external validity concerns.

EXHIBIT 2
Overview of Experimental Design Tax Rates on Investment Gains Associated with Each Period

	Period 1	Period 2	Treatment Introduction	Period 3	Period 4	Period 5	Period 6	Period 7
Immediate Decrease Condition	36%	36%	Tax Decrease to take effect all in the next period	15%	15%	15%	15%	15%
Gradual Decrease Condition	36%	36%	Tax Decrease to be phased in over the next three periods	29%	22%	15%	15%	15%
Immediate Increase Condition	15%	15%	Tax Increase to take effect all in the next period	36%	36%	36%	36%	36%
Gradual Increase Condition	15%	15%	Tax Increase to be phased in over the next three periods	22%	29%	36%	36%	36%

Pre-Tax Change Periods (Periods 1–2)   **Post-Tax Change Periods** (Periods 5–7)

return for the most recent year, while 31.8 percent used some type of paid preparer, and 22.1 percent had their returns prepared by friends or relatives.

The majority of the participants indicated some involvement in the stock market (mean = 3.13 on a scale of no involvement = 1, frequent involvement [6–7 transactions/year] = 5, to extensive involvement [>12 transactions/year] = 9). Only 26 percent indicated no involvement with the stock market. Nearly half (47 percent) indicated they had invested in mutual funds. Thus, most participants appear to have the requisite knowledge for the experimental task.

To provide motivation to exert strong efforts during the experiment, a monetary incentive was provided based on the participant's ability to maximize after-tax returns. To avoid "end-of-game" effects, we informed participants that they would receive a payment based on their after-tax income performance for a randomly selected period relative to the performance of others for the same period. The participant with the highest earnings within each experimental condition for that period received $50, while the other participants received lower amounts, such that payments ranged from $6 to $50 and the average compensation was $10.[13]

[13] Participants may have engaged in a variety of strategies to allocate their funds. In particular, some may have invested all of their funds in the low-risk fund, because it never has a loss. In contrast, standard economic theory would predict that some participants may allocate 100 percent of their funds to the high-risk fund, because it has a much higher expected return than does the lower-risk asset. Very few participants allocated 100 percent of their investment to only one of the funds in every period. Not more than three participants followed this strategy in any group (Immediate Tax Decrease, two participants; Gradual Tax Decrease, one participant; Immediate Tax Increase, two participants; Gradual Tax Increase, two participants; Control Group, three participants). Including these participants in the sample biases against finding support for our predictions because these participants will have zero for the Net Change variable, our dependent variable.

TABLE 1

Demographic Data
(n = 117)

Age	
Mean	25.59
(Standard Dev.)	(4.00)
Number of federal returns filed in the last five years	
Mean	3.68
(Standard Dev.)	(1.71)
Extent to which you completed your own federal tax return[a]	
Mean	4.98
(Standard Dev.)	(3.43)
Form(s) filed in the most recent year	
1040EZ	38.5%
1040A	27.4%
1040	25.6%
Schedule A (itemized deductions)	12.8%
Schedule C (self-employment)	4.3%
Schedule D (capital gains and losses)	10.3%
Schedule E (rental income)	2.6%
Did not file return	8.5%
Who prepared return in the most recent year	
Myself	46.2%
CPA	19.7%
Tax preparation service	11.1%
Other paid preparer	1.0%
Friends/relative	22.1%
Involvement in the stock market[b]	
Mean	3.13
(Standard Dev.)	(1.99)
Previously invested in mutual funds	47.0%

[a] 1 = Relied exclusively on preparers/9 = Exclusively prepared own return.
[b] 1 = No involvement/9 = Extensive involvement (> 12 transactions per year).

We administered the experiment via the Internet so that participants could complete the study when convenient. This administration also allows for control over the order of information and responses. Because monetary rewards are based on the after-tax return of each participant relative to others, the reward system helps to ensure participants complete the study independently. We also reminded participants during the experiment to work on their own.

IV. RESULTS

Hypothesis Test Results

Table 2 provides the mean (standard deviation) of dollars invested in the riskier asset (Mutual Fund B) by period for the four treatment groups and the control group, while Figure 1

TABLE 2

Mean (Standard Deviation) Dollars Invested in the Riskier Asset (Mutual Fund B) by Period

	Period 1	Period 2	Period 3	Period 4	Period 5	Period 6	Period 7
Immediate Tax Decrease							
Mean	$11,113	$10,833	$13,160	$12,460	$12,400	$11,840	$13,000
Standard Deviation	($4,618)	($4,525)	($5,137)	($5,928)	($5,276)	($6,686)	($6,293)
Tax Rate	36%	36%	15%	15%	15%	15%	15%
Gradual Tax Decrease							
Mean	11,968	10,639	12,260	13,652	14,391	15,391	14,739
Standard Deviation	(4,864)	(5,060)	(4,709)	(4,674)	(6,096)	(4,669)	(4,565)
Tax Rate	36%	36%	29%	22%	15%	15%	15%
Immediate Tax Increase							
Mean	12,621	11,316	12,699	12,792	12,941	12,795	12,205
Standard Deviation	(5,390)	(5,451)	(6,098)	(5,912)	(5,884)	(6,024)	(6,439)
Tax Rate	15%	15%	36%	36%	36%	36%	36%
Gradual Tax Increase							
Mean	11,565	9,087	11,478	11,261	9,369	8,957	10,043
Standard Deviation	(5,655)	(5,273)	(5,583)	(5,216)	(6,749)	(6,335)	(6,399)
Tax Rate	15%	15%	22%	29%	36%	36%	36%
Control							
Mean	10,395	9,214	9,855	10,205	10,182	10,955	10,853
Standard Deviation	(5,313)	(6,109)	(6,344)	(6,099)	(6,175)	(6,442)	(5,693)

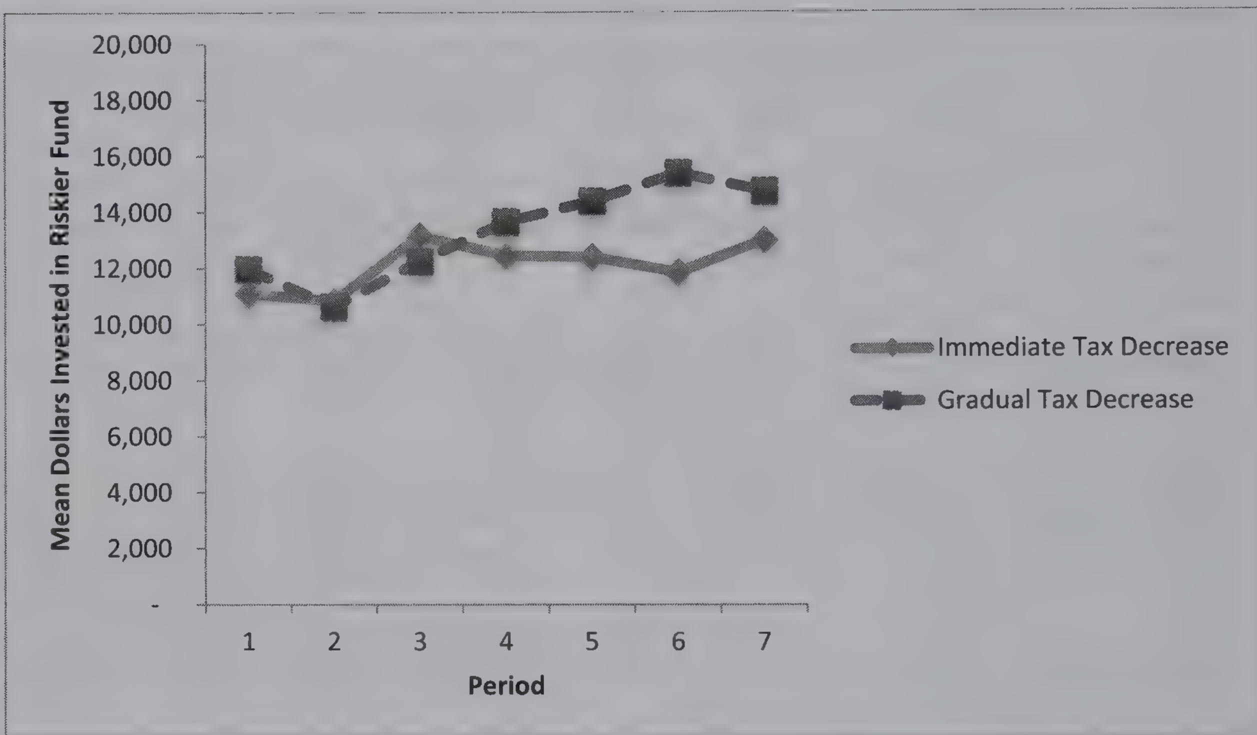
This table presents the mean level of investment in the riskier asset (Mutual Fund B) during each period by treatment group. Note that Period 3 is the initial period of tax change. For the Immediate groups, the tax change is fully implemented in Period 3, and remains at the new tax rate through Period 7. For the Gradual groups, the tax change is fully implemented in Period 5 (after gradual changes in Periods 3, 4, and 5), and remains at the new tax rate through Period 7.

plots the mean investment in the riskier asset for each period for each of the treatment groups. The maximum amount that could be invested in the riskier asset per period is $20,000. As can be seen in Periods 1 and 2, the base periods of no tax change, the mean dollars invested across the conditions range from $9,087 to $12,621. These first two pre-tax change periods provide a baseline for comparison with the effects of the timing and direction of the tax change as measured in the subsequent post-tax change Periods 5–7 when the tax changes have been fully implemented.

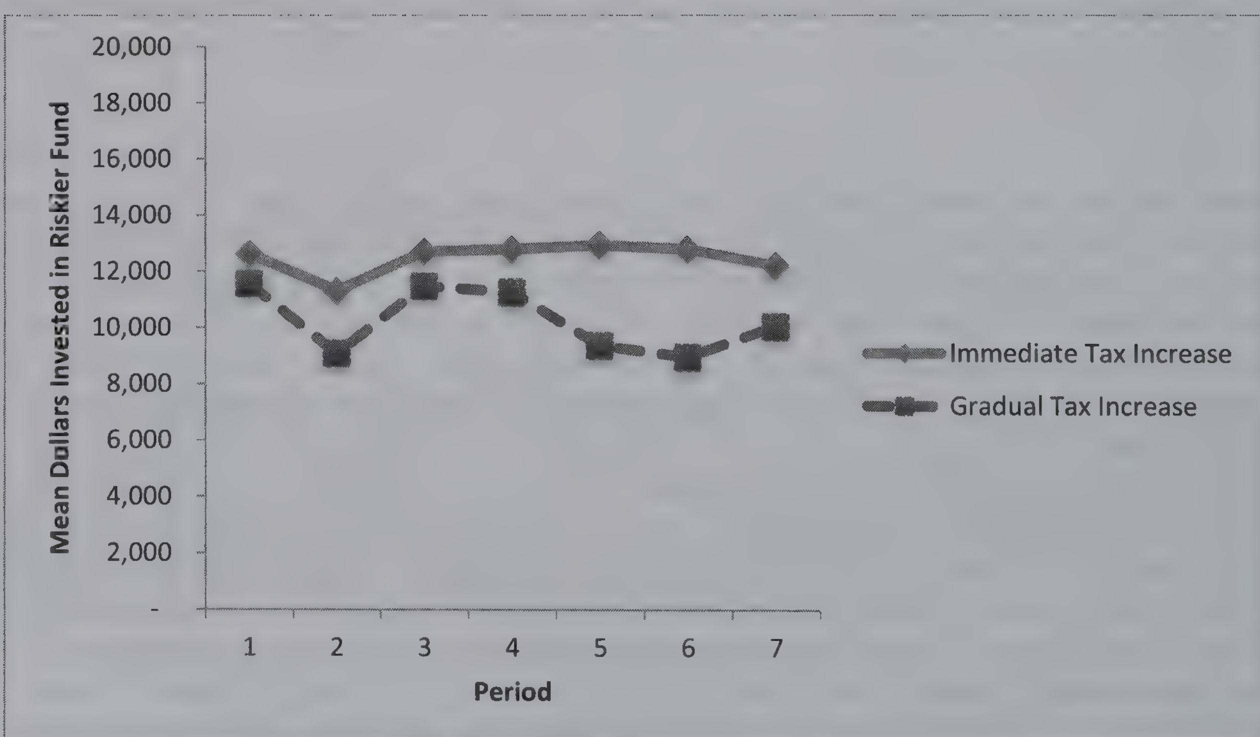
In addition to providing a baseline to measure changes in the investment in the riskier fund, the first two periods also provide an opportunity to test for any differences in the aggregate risk preference profile between the different treatment conditions. While we would expect that the random assignment of participants to the conditions would ensure that the average risk preference would be similar across treatment conditions, testing for differences in the mean amount invested in the riskier fund in Periods 1 and 2 provides additional assurance that the average risk preferences are similar. To test for any differences, we ran two separate ANOVA models for the increase and decrease conditions, with the average investment in Periods 1 and 2 serving as the dependent variable. In the two tax decrease (increase) conditions, the results of this analysis (untabulated) suggest there are no differences in the

FIGURE 1
Mean Dollars Invested in the Riskier Fund by Period

Panel A: Tax Decrease



Panel B: Tax Increase



This figure presents the mean level of dollars invested in the riskier asset (Mutual Fund B) during each period by treatment group.

mean amount invested in the risky fund for these initial periods ($F = 0.072$; $p = 0.79$ and $F = 1.239$; $p = 0.27$ respectively).[14]

For the primary test of our hypotheses, we run a regression model to test our directional expectations. For this analysis, we used the "adjusted" net change in investment as the dependent variable, as described below. In addition to main effects for the gradual or immediate timing of the tax change and the direction of the tax change (increase or decrease), we include an interaction of the two variables to test our research hypotheses. We also include covariates for the number of returns filed in the previous five years (Tax Returns), who prepared the participant's most recent return (Preparer), whether they had previously invested in mutual funds (Mutual Funds), and their work experience (Work Experience). We included these variables to control for variations in the experience of participants in the tax and investment decision contexts as well as in more general contexts. Prior studies have found tax experience (Ayers et al. 1999; Rupert and Wright 1998) and investment experience (Rupert and Wright 1998) to be significant in explaining tax-related and investment decisions.[15]

We calculate the dependent variable as the change in the investment in the riskier asset fund from the pre-tax change periods to the post-tax change periods. The investment in the riskier asset for the pre-tax period is calculated as the average for Periods 1 and 2, while the investment in the riskier asset for the post-tax period is the average investment for Periods 5, 6, and 7.

Because participants' allocations could be influenced by the outcomes of a particular period, in our analyses we use the responses from the control group to gain an understanding of the effects of the outcomes on the level of investment in the absence of any tax changes. The average investment by the control group in the riskier fund for the pre-tax periods was $9,804 and the average investment in the post-tax change periods was $10,663, resulting in an increase of $859. To reflect the impact of the tax changes more accurately, we calculate an "adjusted" change in investment by subtracting this increase from the change in investment for the participants in the treatment groups. Adjusting responses in this manner also provides a better exposition of the results, such that they are directionally consistent with our prediction of a decrease in investment in the tax increase condition, as shown in the last column of Table 3.[16] Table 3 shows the means and standard deviations for the unadjusted pre-tax change periods and post-tax change periods, as well as the net change for these groups, adjusted for the change in investment by the control group.

The results in Table 4 of the regression analysis used to test our research hypotheses show that the interaction between the direction and the timing of the tax change is significant ($t = -2.108$, $p =$

---

[14] We also ran an ANOVA model to compare the average investment in Periods 1 and 2 for all four treatment groups combined. The results of this analysis indicate no differences in the mean amount invested in the riskier fund for the initial periods across the four groups ($F = 0.504$; $p = 0.68$). Given that the tax increase and decrease groups face substantially different tax rates on investment gains in Periods 1 and 2 (36 percent in the tax decrease conditions versus 15 percent in the tax increase conditions), the fact that there are no significant differences in the investment in the riskier portfolio between the tax increase and decrease groups may seem surprising. However, given that the participants are facing a new decision context, it may be that they are approaching the first several periods using an anchoring-and-adjustment type of strategy (Tversky and Kahneman 1974; Carlson 1990). In fact, an examination of the responses for the participants for Periods 1 and 2 suggests that over 40 percent chose to divide their investment approximately 50/50 between the two portfolios. Our study focuses on how participants adjust their investments in response to the changes in the tax system. To examine this question, the key result is that there are no differences in the first two periods between the two tax increase groups and between the two tax decrease groups.

[15] We include these variables in our model to control for potential differences in tax reporting and investment background, and, thus, to reduce noise in the model. However, we ran the model again without these variables and the interaction that tests our hypotheses remains significant at $p = 0.052$.

[16] We also ran the analyses with the unadjusted means and obtain the same results.

TABLE 3

Means and Standard Deviations for the Change in Investment

	Pre-Tax Change in Investment[a]	Post-Tax Change in Investment[b]	Adjusted Net Change in Investment[c]
Immediate Tax Decrease (n = 25)			
Mean	$10,973	$12,413	$581
Standard Deviation	4,129	5,166	4,749
Gradual Tax Decrease (n = 23)			
Mean	11,304	14,841	2,677
Standard Deviation	4,407	4,686	4,078
Immediate Tax Increase (n = 24)			
Mean	11,969	12,647	(181)
Standard Deviation	5,151	5,119	5,274
Gradual Tax Increase (n = 23)			
Mean	10,326	9,457	(1,729)
Standard Deviation	4,958	5,324	2,665

[a] Pre-tax change is calculated as the mean of the investment in the riskier asset (Mutual Fund B) in Periods 1 and 2.
[b] Post-tax change is calculated as the mean of the investment in the riskier asset (Mutual Fund B) in Periods 5, 6, and 7.
[c] Net change adjusted is the amount invested in the riskier asset (Mutual Fund B) in the post-tax change periods minus the amount invested in the riskier asset in the pre-tax change periods. This change is then adjusted by subtracting $859, the increase in the investment in the riskier asset by the control group with no tax implications over these periods. For example, in the immediate tax decrease condition $581 is calculated as ($12,413 – $10,973) – $859.

0.038), supporting the existence of an ordinal interaction.[17] To examine the directional expectations of the subsidiary hypotheses, H1A and H1B, we ran separate regression analyses (untabulated) on each of the Direction of the Tax Change conditions (decrease and increase). For the tax decrease conditions, the additional analysis reveals a significant difference in the change in investment in the riskier asset between the participants in the gradual versus the immediate change conditions ($t = -2.03$, $p = 0.025$, one-tailed). As predicted by the mental accounting and hedonic editing hypothesis (H1A), the participants in the gradual tax decrease condition increased their investment in the riskier fund by a greater amount than the participants in the immediate tax decrease condition. In the gradual decrease condition, the mean net adjusted change in investment was $2,677, while the mean net adjusted change in investment was only $581 for the immediate decrease condition. Figure 2, Panel A shows the change in the investment from the pre-tax change periods to the post-tax change periods for these conditions.

For the tax increase conditions, the supplemental analysis indicates that the difference in the change in the investment in the riskier asset between the gradual versus the immediate timing

[17] In addition to the significant interaction, the results in Table 4 also indicate that the main effect for the Direction of the Tax Change is significant ($t = 3.370$, $p = 0.001$), suggesting, as expected, that the level of investment in the riskier fund is significantly greater for a tax decrease than a tax increase. There is no significant main effect for the Timing of the Tax Change. However, main effects cannot be interpreted unambiguously in the presence of a significant interaction.

TABLE 4

Regression Results
Dependent Variable: Adjusted Net Change in Investment in the Riskier Asset Fund[a]

Variable	Unstandardized Beta Coefficient	t-statistic	p-value
Intercept	−3404.390	−1.390	0.168
Timing of Tax Change[b]	1670.254	1.234	0.220
Direction of Tax Change[c]	4598.935	3.370	0.001
Timing × Direction	−3982.612	−2.108	0.038
Tax Returns[d]	20.133	0.101	0.920
Preparer[e]	−207.555	−0.839	0.404
Mutual Funds[f]	1156.009	1.084	0.281
Work Experience[g]	101.806	0.646	0.520

Model $F = 1.859$ ($p < 0.08$)

Adjusted $R^2 = 0.06$ ($n = 94$)

[a] The dependent variable is Adjusted Net Change in Investment, as shown in the last column of Table 3. This variable is calculated as the difference in the post-tax change periods and pre-tax change periods, adjusted for the change in the investment in the riskier asset (Mutual Fund B) by the control group.
[b] Timing of Tax Change = an independent variable with two levels, where 1 is an immediate change and 0 is a gradual change.
[c] Direction of Tax Change = an independent variable with two levels, where 1 is a tax decrease and 0 is a tax increase.
[d] Tax Returns = a control variable representing the number of returns filed in the previous five years.
[e] Preparer = a control variable representing whether the latest return filed was self-prepared or prepared by a preparer.
[f] Mutual Funds = a control variable representing whether the participants previously invested in mutual funds.
[g] Work Experience = a control variable representing the number of years of work experience.

conditions is marginally significant at $p = 0.055$ ($t = 1.631$, one-tailed), providing support for H1B. Mental accounting predicts that people prefer a large loss versus several smaller losses of the same magnitude (i.e., experience fewer unpleasant feelings). This relative negative affect results in a lower reduction in risk-taking. For this reason, H1B predicts that the participants who experience a tax increase will reduce their investment in the riskier asset more when the tax increase is implemented gradually than immediately. The mean net adjusted change in investment was a decrease of $1,729 in the gradual increase condition, while the mean net adjusted change was a decrease of only $181 in the immediate increase condition.[18] Figure 2, Panel B shows the change in net investment for the two tax increase conditions. In all, the results support hypotheses H1A and H1B.

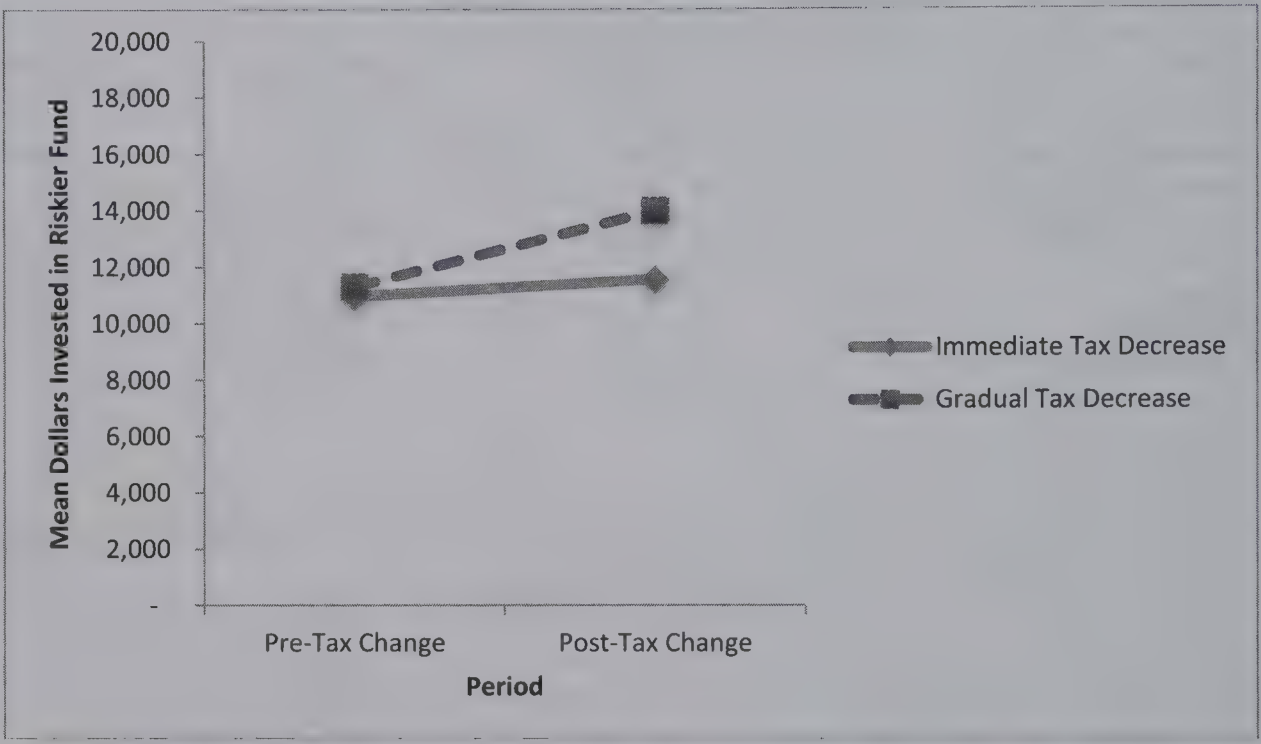
V. CONCLUSIONS AND IMPLICATIONS

Before discussing the implications of the findings, it is important to consider the strengths and limitations of the study. First, we simplified the experimental environment by allowing full deductibility of capital losses. The current limitations on deductibility of losses in the federal income tax system may affect the willingness to invest in riskier assets. Despite this limitation, the use of a controlled setting permitted a relatively strong design, thus, ensuring a high degree of

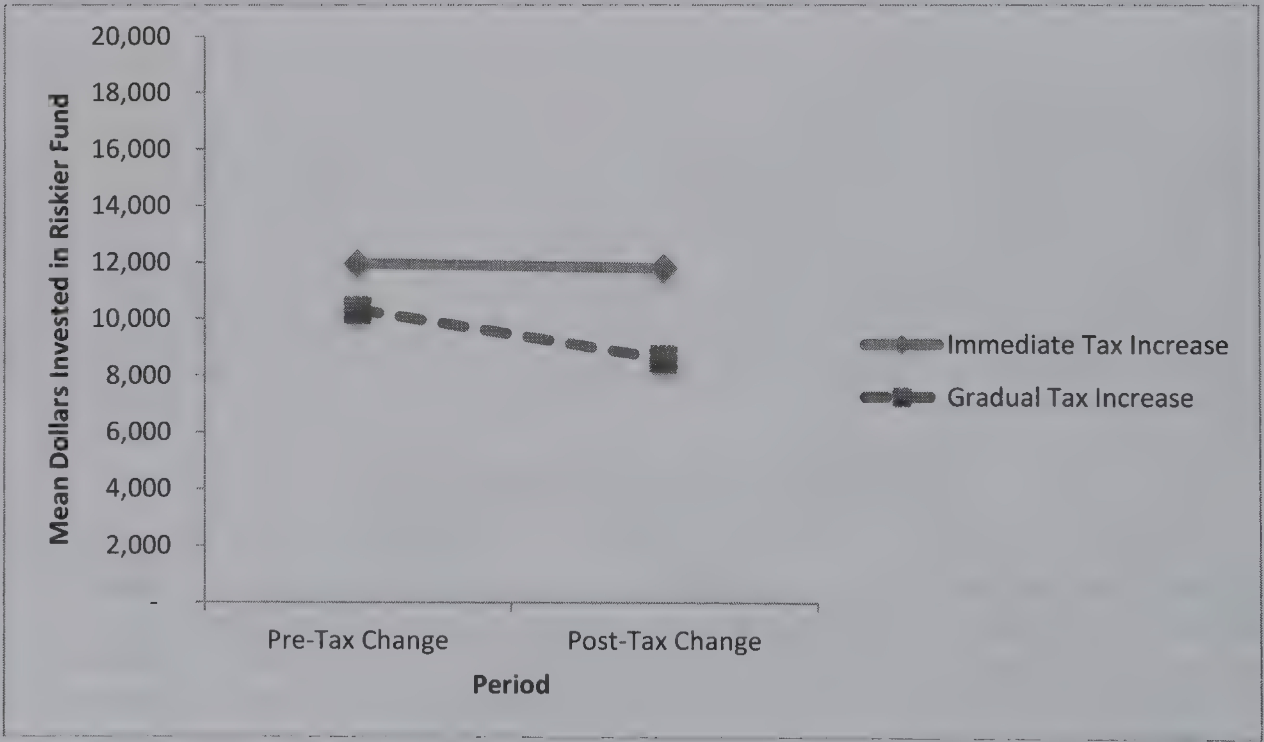
[18] As noted earlier, prospect theory predicts that the value function is steeper for losses than for gains. Therefore, the theory would suggest a larger absolute difference in the change in riskier investments after the tax increase (loss) than after the tax decrease (gain); however, we did not find a significant difference.

FIGURE 2
Mean Dollars Invested in Riskier Fund Pre- versus Post-Tax Change

Panel A: Tax Decrease



Panel B: Tax Increase



This figure presents the mean dollars invested in the riskier asset (Mutual Fund B) during the Pre-Tax Change and Post-Tax Change periods by treatment group. Pre-Tax Change is defined as the mean level of dollars invested in the riskier asset in Periods 1 and 2 (before the tax change). Post-Tax Change is defined as the mean level of dollars invested in the riskier asset in Periods 5, 6, and 7 (after the tax change is fully implemented), adjusted for the control group change (859).

internal validity. Although participants were diverse in age, gender, and income, they consist of M.B.A. students from two universities, whose decisions may not generalize to the entire taxpaying population. To explore the generalizability of our findings, future research should examine these issues while introducing limitations on deductibility of losses and using a sample that is representative of a broader range of taxpayers.

With respect to our design, we use a contextually rich judgment and a judgment and decision-making (JDM) approach rather than one following a strict experimental economics paradigm using abstract terminology. Although we included compensation in an effort to address concerns about whether participants would exert sufficient effort, we do not follow all aspects of an experimental economics approach.[19] Indeed, researchers (Davis and Swenson 1988; Alm 1991; Wartick et al. 1997; Haynes and Kachelmeier 1998) argue for the exclusion of explicit terms, such as "taxes," to avoid conveying unintended information, such as overtones of ethical and/or authoritative considerations that are not part of the economic model being tested, which can result in a loss of experimental control, particularly in studies examining tax compliance. However, our goal was to explicitly examine taxpayers' responses to the timing of tax changes. Given this research objective, we were willing to make the trade-off to obtain a more real-world setting and generalizability, similar to that employed in other tax studies of prospect theory (Schadewald 1989; Schepanski and Kelsey 1990; White et al. 1993; Ayers et al. 1999). Further, our study did not examine an issue of tax compliance, in which the use of tax terminology may have led to unintended role-playing. Using our current design with performance-contingent compensation provides us with a blend of JDM and economic designs that help address the limitations of both.

Over the past several decades Congress has frequently used the federal income tax system as a mechanism for influencing the behavior of taxpayers. The context for the current study is investments in risky assets. Congress has often hoped to encourage such investments, including those in start-up companies. One of the key means for promoting this type of investment has been through preferential treatment of capital gains and losses. As a result, these provisions undergo frequent changes by Congress.

In implementing these changes, Congress has utilized a number of different timing approaches, implementing tax rate changes immediately in some cases, and over a number of years in other cases. In the present study, we examine whether the immediate or gradual timing of the change and the direction of the change as a tax increase or decrease influence the decision to invest in risky assets. The results of our experiment indicate that the timing of the tax change has a significant effect on the funds invested in the riskier fund when the taxpayer faces a tax decrease or tax increase.

The findings for the tax decrease conditions are consistent with mental accounting. We find that taxpayers exhibit a greater preference for riskier investments when the tax decrease is implemented gradually rather than immediately, consistent with the preference to segregate gains in the form of tax rate decreases over multiple periods in mental accounting and hedonic editing.

The results for the tax increase conditions also provide support that taxpayers prefer to integrate losses by facing unpleasant events at one time, as mental accounting and hedonic editing would predict. Those participants who experienced an immediate tax increase reduced their investment in the riskier fund by a smaller amount than those who experienced a gradual tax increase. These findings are consistent with recent studies on the role of affect on risk-taking, which

[19] In a commentary on using laboratory experiments to evaluate accounting policy issues, Kachelmeier and King (2002, 222) state that "without performance-contingent compensation, it is unknown whether participants in these and related individual-judgment studies exert as much information-processing effort in the experimental setting as in an environment with explicit incentives."

American Accounting Association

have shown that pleasant (unpleasant) feelings lead to increased (decreased) risk-taking with gains (losses) (Seo et al. 2010; Au et al. 2003; Isen and Patrick 1983).

The findings of the present study offer several possibilities for future research. First, our study is set in the context of changes in capital gains taxes and risky investment. An important issue is the extent to which the results hold for other tax changes. Our hypotheses are grounded in the theory developed in mental accounting and the hedonic editing hypothesis (Thaler 1985; Thaler and Johnson 1990), which focuses on risky choices. Thus, we are reluctant to claim that our hypotheses will apply for any tax rate change. For instance, while the timing and direction of a change in the marginal income tax rate also appear to affect taxpayers' utility in a manner consistent with predictions of mental accounting regarding, for instance, the choice between work and leisure, empirical evidence is needed to support this presumption. Second, in this study, we examined substantial changes in the tax rate (i.e., an increase from 15 percent to 36 percent or a decrease from 36 percent to 15 percent). Future research might examine whether taxpayers respond in a similar way to more modest changes in the tax rate on capital gains.

In our study we focus on investment decisions for seven periods and do not investigate whether the level of investment going forward remains at that relative level on a permanent basis, an avenue for future research. A final potential topic for future research is an examination of various methods of implementing alternative tax changes beyond tax rate changes. While changes in tax rates are perhaps the most explicit way that Congress can increase or decrease taxes, Congress may also change a taxpayer's level of taxation by changing exclusions, deductions, or credits. Future research could examine whether alternative means of implementing these types of tax changes have similar effects to tax rate changes.

Over the last decade very few tax studies have employed experimental designs. However, an experimental methodology is an excellent tool with which to examine tax issues, since it provides the researcher with the ability to examine economic and psychological theories in a controlled environment, while discerning cause and effect. Unlike experimental data, naturally occurring or observed data from capital markets or the IRS are subject to many influences and factors that are not held constant. In addition, natural data cannot be assembled to examine tax issues *ex ante*, such as the one in our study, since proposed tax provisions do not exist. Given these limitations, employment of an experimental method is fruitful for future tax research, particularly when examining potential tax policy changes. Along these lines, a review by Alm (2010) highlights the recent increase in the use of laboratory experiments to examine various applications of behavioral economics to the specific area of public economics. In particular, behavioral public economics can examine issues of public goods, tax compliance, and individual responses to taxes.

The use of experimental methodologies to examine tax policy research questions is, thus, quite fertile. As Congress currently meets to discuss tax reform, many issues are on the table, such as a new tax system that may entail a value-added tax (VAT). In addition, these issues permeate to state jurisdictions, which currently debate ways to raise additional revenue to cover budget constraints. As these policy discussions continue, many opportunities become available to examine taxpayers' responses to such systems.

REFERENCES

Allingham, M. G., and A. Sandmo. 1972. Income tax evasion: A theoretical analysis. *Journal of Public Economics* 1: 323–338.

Alm, J. 1991. A perspective on the experimental analysis of taxpayer reporting. *The Accounting Review* 66 (July): 577–593.

Alm, J. 2010. Testing behavioral public economics theories in the laboratory. *National Tax Journal* 63 (4): 635–658.

Anderson, S. E., and J. B. Butler. 1997. Experimental evidence on the effects of tax incentives on risky security prices. *The Journal of the American Taxation Association* 19 (1): 58–76.

Arkes, H. R., D. Hirshleifer, D. Jiang, and S. Lim. 2008. Reference point adaptation: Tests in the domain of security trading. *Organizational Behavior and Human Decision Processes* 105: 66–81.

Au, K., F. Chan, D. Wang, and I. Vertinksy. 2003. Mood in foreign exchange trading: Cognitive processes and performance. *Organizational Behavior and Human Decision Processes* 91: 322–338.

Ayers, B. C., S. J. Kachelmeier, and J. R. Robinson. 1999. Why do people give interest-free loans to the government? An experimental study of interim tax payments. *The Journal of the American Taxation Association* 21 (2): 55–74.

Burman, L. 1999. *The Labyrinth of Capital Gains Tax Policy: A Guide for the Perplexed.* Washington, DC: Brookings Institute Press.

Carlson, B. W. 1990. Anchoring and adjustment in judgments under risk. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16 (4): 665–676.

Chatterjee, S., T. B. Heath, and J. Min. 2009. The susceptibility of mental accounting principles to evaluation mode effects. *Journal of Behavioral Decision Making* 22 (2): 120–137.

Cowell, F. A. 1990. *Cheating the Government.* Cambridge, MA: The MIT Press.

Cowley, E. 2008. The perils of hedonic editing. *Journal of Consumer Research* 35: 71–84.

Davis, J. S., and C. Swenson. 1988. The role of experimental economics in tax policy research. *The Journal of the American Taxation Association* 10 (Fall): 40–59.

Domar, E. D., and R. A. Musgrave. 1944. Proportional income taxation and risk-taking. *Quarterly Journal of Economics* 58 (3): 388–422.

Dunn, J. 1998. Dunn comments on small business investment initiative. *Tax Notes* (October 12): 231.

Fellingham, J. C., and M. A. Wolfson. 1978. The effects of alternative income tax structures on risk-taking in capital markets. *National Tax Journal* 31 (4): 339–347.

Goldberg, D. S. 1994. Tax subsidies: One-time v. periodic, an economic analysis of the tax policy alternatives. *Tax Law Review* 49: 305–347.

Graetz, M. J. 1977. Legal transitions: The case of retroactivity in income tax revision. *University of Pennsylvania Law Review* 126: 47–87.

Haynes, C. M., and S. J. Kachelmeier. 1998. The effects of accounting contexts on accounting decisions: A synthesis of cognitive and economic perspectives in accounting experimentation. *Journal of Accounting Literature* 17: 97–136.

Isen, A. M., and R. Patrick. 1983. The effect of positive feelings on risk-taking: When the chips are down. *Organizational Behavior and Human Decision Processes* 31: 194–202.

Kachelmeier, S. J., and R. R. King. 2002. Using laboratory experiments to evaluate accounting policy issues. *Accounting Horizons* 13 (3): 219–232.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47: 263–291.

Kahneman, D., and A. Tversky. 1984. Choices, values, and frames. *The American Psychologist* 39: 341–350.

Kaplow, L. 1986. An economic analysis of legal transitions. *Harvard Law Review* 99: 509–617.

King, R. R., and D. E. Wallin. 1990. Individual risk-taking and income taxes: An experimental examination. *The Journal of the American Taxation Association* 12 (1): 26–38.

Lerner, J. 2002. Symmetrical treatment: Capital gains and capital losses. *Tax Notes* (December 9): 1361–1362.

Lim, S. S. 2006. Do investors integrate losses and segregate gains? Mental accounting and investor trading decisions. *Journal of Business* 79 (5): 2539–2573.

Linville, P. W., and G. W. Fischer. 1991. Preferences for separating or combining events. *Attitudes and Social Cognition* 60 (1): 5–23.

Logue, K. D. 1996. Tax transitions, opportunistic retroactivity, and the benefits of government precommitment. *Michigan Law Review* 94 (5): 1229.

Meade, J. A. 1990. The impact of different capital gains regimes on the lock-in effect and new risky investment decisions. *The Accounting Review* 65 (2): 406–431.

National Taxpayer Advocate. 2008. *Annual Report to Congress.* Available at: http://www.irs.gov/advocate

Rupert, T. J., and A. M. Wright. 1998. The use of marginal tax rates in decision making: The impact of tax rate visibility. *The Journal of the American Taxation Association* 20 (2): 83–99.

Schadewald, M. S. 1989. Reference point effects in taxpayer decision making. *The Journal of the American Taxation Association* (Spring): 69–84.

Schepanski, A., and D. Kelsey. 1990. Testing for framing effects in taxpayer compliance decisions. *The Journal of the American Taxation Association* (Fall): 60–77.

Schneider, D. 1980. The effects of progressive and proportional income taxation on risk-taking. *National Tax Journal* 33 (1): 67–75.

Seo, M., B. Goldfarb, and L. F. Barrett. 2010. Affect and the framing effect within individuals over time: Risk taking in a dynamic investment simulation. *Academy of Management Journal* 53: 411–431.

Sikes, S., and R. Verrecchia. 2012. Capital gains taxes and expected rates of return. *The Accounting Review* 87 (3): 1067–1086.

Stiglitz, J. E. 1969. The effects of income, wealth, and capital gains taxation on risk-taking. *Quarterly Journal of Economics* 83: 263–283.

Tax Notes Today. 1997. Hatch bill would reduce tax rate on capital gains. (Doc. 97-2598). *97 TNT*: 21–65.

Thaler, R. 1985. Mental accounting and consumer choice. *Marketing Science* 4 (3): 199–215.

Thaler, R. 1999. Mental accounting matters. *Journal of Behavioral Decision Making* 12: 183–206.

Thaler, R., and E. J. Johnson. 1990. Gambling with the money and trying to break even: The effects of prior outcomes on risky choice. *Management Science* 36 (6): 643–660.

Tversky, A., and D. Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185: 1124–1131.

Tversky, A., and D. Kahneman. 1981. The framing of decisions and the rationality of choice. *Science* 211: 453–458.

Von Neumann, J., and O. Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Wartick, M. L., S. A. Madeo, and C. C. Vines. 1997. Reward dominance in tax-reporting experiments: The role of context. *The Journal of the American Taxation Association* 21 (1): 20–31.

White, R. A., P. D. Harrison, and A. Harrell. 1993. The impact of income tax withholding on taxpayer compliance: Further empirical evidence. *The Journal of the American Taxation Association* 15 (2): 63–78.

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 521–552

American Accounting Association
DOI: 10.2308/accr-50322

The Contagion Effect of Low-Quality Audits

Jere R. Francis
University of Missouri–Columbia

Paul N. Michas
The University of Arizona

ABSTRACT: We investigate if the existence of low-quality audits in an auditor office indicates the presence of a "contagion effect" on the quality of other (concurrent) audits conducted by the office. A low-quality audit is defined as the presence of one or more clients with overstated earnings that were subsequently corrected by a downward restatement. We document that the quality of audited earnings (abnormal accruals) is lower for clients in these office-years (when the misreporting occurred) compared to a control sample of office-years with no restatements. This effect lasts for up to five subsequent years, indicating that audit firms do not immediately rectify the problems that caused contagion. We also find that an office-year with client misreporting is likely to have subsequent (new) client restatements over the next five fiscal years. Overall, the evidence suggests that certain auditor offices have systematic audit-quality problems and that these problems persist over time.

Keywords: *audit quality; auditor offices; contagion.*

Data Availability: *All data are publicly available.*

I. INTRODUCTION

We investigate if the existence of at least one low-quality audit in an auditor office location indicates a more systematic problem in office-level audit quality for publicly traded clients. The term "audit failure" is used to refer to audit engagements in which there is a downward restatement of previously audited client earnings.[1] A "contagion" of

We thank Mike Ettredge (editor), the two reviewers, and John Harry Evans (senior editor) for their constructive comments and guidance. We also appreciate comments on the paper when presented at the 2011 Annual Meetings of the American Accounting Association, European Accounting Association, and International Symposium on Audit Research, and workshops at The University of Arizona, University of Illinois, Indiana University, The University of Melbourne, University of Missouri, Nanyang Technological University (Singapore), The Ohio State University, Texas A&M University, and The University of Texas at Dallas.

We thank the Center for Audit Quality for providing funding for this research project. However, the views expressed in this article and its content are those of the authors alone and not those of the Center for Audit Quality.

Editor's note: Accepted by Michael Ettredge.

Submitted: September 2011
Accepted: October 2012
Published Online: October 2012

[1] The term "audit failure" can be defined more narrowly, such as court judgments or SEC enforcement actions against auditors, although the Panel on Audit Effectiveness (2000, paras. 1.6 and 3.26) notes that a restatement is strongly suggestive that the audit of the originally issued financial statements was of unacceptably low quality. We also believe the use of accounting restatements can provide insight on a much wider range of potentially low-quality audits than a narrower definition of audit failures.

low-quality audits could occur in an auditor office location due to office-specific characteristics including personnel and quality-control procedures. Gleason et al. (2008, footnote 8) define contagion as occurring "when an adverse event at one firm also conveys negative information about . . . other firms." In our test setting, contagion occurs if the presence of one low-quality audit in an engagement office conveys negative information about the quality of other concurrent audits conducted by the office.

Prior research provides evidence that differences in characteristics across offices of accounting firms are an important determinant of audit quality, and that differences in audit quality can exist even within the same audit firm, depending on office-level characteristics. For example, Francis and Yu (2009) and Choi et al. (2010) show that audit quality is higher in larger Big 4 auditor offices. Research also shows that industry expertise within an office is positively associated with engagement-specific audit quality and audit pricing (Francis et al. 2005; Reichelt and Wang 2010). This research highlights the importance of investigating auditor office-level characteristics and their effects on audit quality, and is consistent with the view that offices are the primary decision-making units in accounting firms (Wallman 1996).

Our results show that offices with an audit failure are more likely to have additional (new) audit failures in the subsequent five years, suggesting a longitudinal contagion of audit failures over time. We also find that concurrent clients in offices with audit failures have a higher level of abnormal accruals compared to offices with zero audit failures, which is suggestive of lower earnings quality (Francis et al. 1999a; Francis and Yu 2009; Reichelt and Wang 2010). These results hold for all but the largest quartile (top 25 percent) of auditor office size. Last, in a separate analysis of Big 4 offices we document that the contagion effect is mitigated for the smallest 75 percent of offices when a large portion of audits are in those industries in which the office is the city-level industry leader. Thus, there is interplay between office size and the office's level of industry expertise, and their effect on audit quality.

Our study can help multiple parties in assessing office-specific audit quality. Regulators such as the PCAOB can focus inspections on auditor office locations that are more likely to be problematic. Audit standard-setters may formulate auditing standards to better address audit-quality problems at the office level, and audit firms can allocate their resources more effectively to improve quality-control procedures in those offices more likely to conduct persistently low-quality audits.[2] Finally, investors may be able to use the results to assess current earnings quality based on the auditor office's prior history of audit failures.

The paper proceeds as follows. We develop our hypotheses in Section II, and Section III presents the sample, research design, and descriptive statistics. Section IV reports the results, and the paper concludes in Section V.

II. BACKGROUND AND HYPOTHESES

Background

Prior research argues that offices are the primary decision-making units in accounting firms (Francis et al. 1999b; Francis and Yu 2009; Wallman 1996). However, the extant literature that investigates the determinants of audit quality at the auditor office-level is relatively scant. Francis and Yu (2009), Choi et al. (2010), and Francis et al. (2012) are the only studies that currently

---

[2] While audit failure information has obviously been available to national offices in the past, ours is the first study that empirically investigates whether any party, including the national office of audit firms, can infer something systematic about an office's audit quality by identifying a specific audit failure using publicly available information.

provide a way to distinguish *overall* audit quality at the office level by providing evidence that the size of an auditor office is positively associated with audit quality. While these studies attempt to look into the "black box" of auditor offices to investigate office-level characteristics associated with differential audit quality, office size is a somewhat crude tool that may not be as useful to outsiders as a measure that is more specific. Further, given that auditor office size is likely to be very stable from year to year, this measure is not able to discern yearly variations in office-level audit quality. Given the relatively high amount of turnover within audit firms (Hiltebeitel et al. 2000), a measure that provides an indication of overall audit quality within an office in a particular year is likely to be more useful compared to office size alone. We show how our measure can be used in conjunction with auditor office size, thus providing an important contribution to the literature as well as a more refined way to assess office-level audit quality. Our study examines audit failures separately for offices of Big 4 and non-Big 4 accounting firms. This is important because we know relatively little about non-Big 4 firms, yet they conduct audits for about 30 percent of publicly traded companies and their market share has grown since the collapse of Arthur Andersen in 2002.

Prior research also documents differences in engagement-specific audit quality based on an auditor office's industry expertise (e.g., Reichelt and Wang 2010). However, while an auditor office is classified as an expert in particular industries, that office will typically audit many clients outside of its areas of industry expertise. In other words, the unit of analysis in these studies is engagement-specific industry expertise, not a more general office-wide measure of auditor expertise. In contrast, we compare the quality of all audits in offices where audit failures are identified, with all audits in those offices where no audit failures are identified. Therefore, we are investigating inter-office variation in an office characteristic—the presence or absence of an audit failure—instead of variation in engagement-specific industry expertise. However, we control for engagement-specific industry expertise in the primary tests, and we also conduct an additional analysis to determine if the overall use of such expertise with an office mitigates office-level contagion.

Hypotheses Development

An audit failure in an auditor office-year may indicate one of two possibilities. First, it may indicate that a one-off audit engagement was of low quality for engagement-specific or idiosyncratic reasons. The second possibility, and the one that we investigate, is that one audit failure may *reveal* a more systematic problem in an office due to general characteristics of the office. We term this a *contagion*. Specifically, it is possible that general characteristics of office-level personnel, including an office's level of auditor expertise or the lack of office-level, quality-control procedures, led to the specific audit failure, as well as other low-quality audits. If this is the case, then audit failures may provide useful information about the quality of concurrent audits performed in the office.[3]

We test whether a contagion effect exists at the office level based on the identification of one or more audit failures as having occurred in a specific auditor office in a given fiscal year. This requires us to identify and measure audit failures. Palmrose and Scholz (2004) and Kinney et al. (2004) argue that a material restatement of originally audited financial statements is strongly suggestive that the audit of the original, misstated financial statements was of low quality. This view is reinforced by the Panel on Audit Effectiveness (2000, para. 1.6), which says that "Restatements also raise the question, 'Where were the auditors?'" The report goes on to suggest

[3] For example, Krishnan (2005) analyzes office-level audit quality and finds that the clients of the Houston office of Arthur Andersen, which audited Enron, exhibit less timely reporting of bad news compared to a sample of Houston-based clients audited by other Big 6 audit firms, as well as clients of Andersen's Atlanta office, in the same year as the Enron audit failure.

"Restatements of previously audited financial statements raise questions about whether the system that provides assurances about both the *quality of audits* and the reliability of financial reports is operating effectively" (para. 3.26; emphasis added).

A company may have accounting restatements for various reasons. Plumlee and Yohn (2010) examine 3,744 restatements from 2003–2006 and identify four main causes: the majority (57 percent) of restatements are caused by internal company error followed by characteristics of accounting standards (37 percent), which includes complexity, lack of clarity in the standard, and the need to use judgment in applying the standard. The remaining restatements are due to fraud (3 percent) and transaction complexity (3 percent). We believe that a company's external auditor bears some responsibility for allowing a company to issue financial statements that are misstated due to any of these four causes because auditors have "a responsibility to plan and perform the audit to obtain reasonable assurance about whether the financial statements are free of material misstatement, whether caused by error or fraud" (SAS No. 1, AICPA 1972). Consequently, a high-quality audit should, *ceteris paribus*, detect misstatements due to any of the above reasons at a higher rate compared to a low-quality audit, including the professional judgment required to deal with complexity and the interpretation and implementation of accounting standards. We conclude that the presence of an accounting restatement is indicative that a relatively low-quality audit was performed when the misstated financial statements were originally issued.

We predict that in auditor office-years where at least one client misreports (as evidenced by the subsequent downward restatement of earnings), audit quality is lower on average for other clients audited by that office in the same year. The first hypothesis in alternative form is that the presence of one audit failure reveals a contagion effect on the quality of concurrent audits:

H1: The existence of an audit failure in an auditor office is indicative of a contagion effect that reveals the presence of other concurrent low-quality audits in the office.

Francis and Yu (2009), Choi et al. (2010), and Francis et al. (2012) find that audited earnings are of higher-quality for clients in larger Big 4 auditor offices compared to smaller offices. If larger offices perform higher-quality audits, then we would expect to observe less contagion in larger offices. Francis and Yu (2009) attribute the office-size effect to larger offices possessing more in-house experience with public companies, and therefore greater human capital in the office, while Choi et al. (2010) attribute the result to larger offices being subject to lower economic dependence on any one client. In the context of a contagion effect, these arguments suggest that in large offices an audit failure (on a specific engagement) is more likely to be idiosyncratic rather than symptomatic of widespread problems, and the second hypothesis in alternative form is:

H2: There is less contagion in large Big 4 offices than in small Big 4 offices.

While H2 specially focuses on Big 4 firms, for completeness we also compare large and small offices of non-Big 4 accounting firms.

Finally, the literature on auditor industry expertise indicates that office-specific industry expertise is an important determinant of engagement-level audit quality (Reichelt and Wang 2010). However, audits at the office level are conducted for clients that operate within the office's areas of industry expertise as well as in other industries. Furthermore, the number of audits for which the office is an industry expert, as a percentage of the total number of audits in the office, likely varies across offices. In offices where the vast majority of audits are in the office's areas of industry expertise, engagement personnel are more able to apply their industry-specific knowledge and human capital to the office's overall client portfolio, which should result in high-quality audits. In contrast, in offices where relatively few audits are conducted within the office's areas of industry expertise, audit personnel make less use of their industry-specific knowledge. This leads to the third hypothesis stated in alternative form:

H3: There is less contagion in a Big 4 office where relatively more audits are conducted in the office's areas of industry expertise, compared to Big 4 offices where relatively fewer audits are conducted in the office's areas of industry expertise.

We do not test H3 for non-Big 4 auditors as prior literature considers only Big 4 offices to be city-level industry experts (Francis et al. 2005; Reichelt and Wang 2010).

We test for a contagion effect by comparing the quality of clients' audited earnings in those office-years with an audit failure (treatment sample), with the quality of clients' earnings in office-years with *no* audit failures (control sample). Earnings are jointly produced by the client and the auditor (Antle and Nalebuff 1991). Clients are responsible for preparing the financial statements in accordance with GAAP (Generally Accepted Accounting Principles), and the auditor's role is to enforce compliance with GAAP. The research design linking statistical properties of earnings with audit characteristics is described by Francis (2011):

$$\textit{Audited Earnings Quality} = f(\textit{Audit Attributes} + \textit{Controls}),$$

where earnings quality is measured by cross-sectional variation in statistical properties of audited earnings. In this equation, audited earnings quality is a function of specific audit attributes such as the type of auditor (Big 4 or non-Big 4), auditor office characteristics, or engagement-specific factors. Controls refer to client-level controls and other attributes of audit firms (other than the test variable) that are likely to affect the specific measure of audited earnings quality used in a given analysis. It is important to emphasize that earnings quality metrics are not a direct measure of audit quality. Rather, given that earnings are jointly produced by clients and auditors, cross-sectional differences in the statistical properties of audited earnings suggest that there are differences in the underlying quality of audits, based on systematic auditor characteristics.

Following prior research, we test if auditor characteristics are associated with abnormal accruals (Becker et al. 1998; Francis and Yu 2009; Frankel et al. 2002; Reichelt and Wang 2010). Earnings are assumed to be of higher-quality when abnormal accruals are smaller in magnitude, *ceteris paribus*, and the audit attribute we test is whether the engagement office has an audit failure as evidenced by a subsequent downward earnings restatement. A contagion effect is evidenced if earnings quality is lower on average (larger abnormal accruals) for clients in offices with an audit failure, compared to clients in offices with no audit failures.

III. RESEARCH DESIGN AND DESCRIPTIVE STATISTICS

Sample

As described in more detail in the next subsection, an audit failure occurs in an office when there is a downward restatement of net income by a client subsequent to the statutory audit. The year of the audit failure is the year in which the misstated earnings were *originally* reported. We use the Audit Analytics database to identify restatements and the original filing year for which the financial statements were subsequently restated. We use the Compustat Unrestated U.S. Quarterly Data File to obtain originally released as well as subsequently restated accounting data in order to identify the yearly restatement amount, if any.[4] This database provides originally reported quarterly

[4] The Audit Analytics database provides information about a restatement "period" for each firm, which can be one year or more than one year in length. Further, it identifies the effect on net income, if any, only for the entire restatement period, not for each individual year. Conversely, the Compustat Unrestated U.S. Quarterly file identifies the net income effect for each quarter of each fiscal year for each firm. Studies that use the Compustat Unrestated Quarterly file to obtain originally released accounting data include Bronson et al. (2011), Price et al. (2011), and Comprix et al. (2012), among others.

financial statement data, including net income, and many of the data items available in the Compustat Fundamentals Quarterly and Annual Files. Finally, we limit the restatements to those that cross-map to the Audit Analytics restatements database. The reason is that Audit Analytics restricts its database to accounting restatements related to accounting errors, fraud, and GAAP misapplications. In contrast, there are additional restatements in Compustat due to GAAP changes as well as entity changes due to mergers and acquisitions.

The advantage of using the Compustat database (in conjunction with Audit Analytics) is that it readily identifies the yearly dollar amount of accounting restatements. The Compustat Unrestated Quarterly file also includes the originally released as well as the most current restated values for each data item. For companies for which no restatement took place, the data value is exactly the same in both the unrestated and restated item columns. A company's annual earnings (both unrestated and restated) is computed by summing the four quarters of the fiscal year.[5]

The sample period begins in the year 2000, the first year data on the specific auditor office location is available in the Audit Analytics database. We cut off the sample in 2008 because Cheffers et al. (2010) show that the average time lag between the original financial statement release and a restatement in the years 2005 to 2007 is about 700 days, or roughly two years. Therefore, cutting off the sample in 2008 provides confidence that we are correctly classifying the vast majority of restating and non-restating companies.

Table 1, Panel A summarizes the sample selection. There are 87,890 annual firm-year observations in the Compustat Unrestated Quarterly data file for the years 2000 through 2008 with non-missing assets or income. We delete 25,019 financial and utility companies due to the specific operating and accounting characteristics of these firms. Central Index Key (CIK) numbers are used to merge accounting data with auditor office location information drawn from Audit Analytics and are missing for 7,081 observations, and the specific auditor office location data are missing in Audit Analytics for 11,637 observations. There are 2,260 observations that have a downward restatement of net income, and these are used to measure audit failures in specific office-years as described below. In the contagion tests, we delete these 2,260 observations because we test if a contagion effect occurs for the other concurrent clients of the office.[6] Finally, we delete 19,267 observations due to missing information necessary to compute firm-level accounting variables, including abnormal accruals and stock price-based variables. The final sample is comprised of 22,626 firm-year observations for 4,765 unique companies from 2000 through 2008. Table 1, Panel B indicates there are 2,475 Big 4 office-years in the sample with an average of 275 unique offices per year, while there are 1,997 non-Big 4 offices years with an average of 222 unique offices per year. Untabulated results show that Big 4 (non-Big 4) offices have an average of 15 (8) clients, and the largest Big 4 (non-Big 4) office has 567 (408) clients.

Office-Level Audit Failures

Our measure of an audit failure involves identifying whether one or more clients of a specific auditor office in a given year subsequently have a downward restatement of net income. For these offices we calculate the percentage restatement of a company's annual net income by measuring the dollar value difference in net income between the originally released financial information and the

[5] A small percentage of restatements identified in Audit Analytics are communicated to the SEC before the company's fiscal year-end. Specifically, out of a total of 16,175 restatement-years identified by Audit Analytics, only 628, or 3.88 percent, are reported to the SEC before the company's fiscal year-end date. We do not consider these to be audit failures when calculating our main variable of interest, *AUD_FAIL_X*.

[6] We retain the 1,313 upward restatements in the control sample as we do not consider these to be audit failures. However, all results are very similar if we delete these upward restatements from the control sample.

TABLE 1

Sample

Panel A: Sample Selection

	n
Observations available in the Compustat Unrestated Quarterly Data File from the years 2000–2008 with non-missing assets or income	87,890
Less:	
Financial and Utility Companies (SIC 4400–4999 and 6000–6999)	(25,019)
Observations with missing CIK number to merge with Audit Analytics	(7,081)
Observations with missing auditor location data in Audit Analytics	(11,637)
Observations with a downward restatement of net income	(2,260)
Observations with missing data necessary to calculate firm-level variables	(19,267)
Final Sample	22,626

Panel B: Number of Unique Auditor Office Locations by Year

Year	Non-Big 4	Big 4	Total
2000	99	317	416
2001	161	343	504
2002	196	292	488
2003	235	272	507
2004	272	263	535
2005	294	260	554
2006	283	249	532
2007	226	244	470
2008	231	235	466
Total Office-Years	1,997	2,475	4,472
Mean Number of Offices	222	275	497

(continued on next page)

TABLE 1 (continued)

Panel C: Classification of Office-Years with Audit Failures

Number of Offices With At Least One Audit Failure	Restatement Threshold (X): > 0% Non-Big 4	> 0% Big 4	> 10% Non-Big 4	> 10% Big 4
$AUD\_FAIL\_X = 0$	1,684	1,599	1,684	1,599
$AUD\_FAIL\_X = 1$	313	876	199	515
Total Office-Years	1,997	2,475	1,883	2,114

(continued on next page)

TABLE 1 (continued)

Panel D: Office-Level Descriptive Statistics

	Non-Big 4				Big 4			
	AUD_FAIL_0 = 0 n = 1,684		*AUD_FAIL_0* = 1 n = 313		*AUD_FAIL_0* = 0 n = 1,599		*AUD_FAIL_0* = 1 n = 876	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
OFFICE_SIZE	599	164	1,702***	706***	7,491	2,505	2,290***	11,602***
M_RISK_PORT	1,979	1.055	2.551***	1.944***	−0.443	−0.635	−0.472	−0.538
M_CITY_IND_EXP	NA	NA	NA	NA	0.635	0.667	0.520***	0.500***
M_NAT_IND_EXP	0.005	0	0.006	0***	0.198	0	0.227**	0.119***
M_INFLUENCE	0.794	0.832	0.274***	0.167***	0.542	0.321	0.113***	0.052***
M_SIZE	2.690	2.758	2.667	2.685	5.991	5.997	5.823***	5.819***
M_LAG_TOT_ACC	−0.169	−0.071	−0.225***	−0.097***	−0.065	−0.054	−0.062	−0.057**
M_CFO	−0.257	−0.043	−0.342*	−0.094***	0.057	0.076	0.057	0.071*
M_CFO_VOL	0.270	0.162	0.306**	0.193***	0.101	0.078	0.096	0.082**
M_SALES_GROWTH	0.124	0.052	0.127	0.070	0.102	0.011	0.086**	0.076
M_SALES_VOL	0.071	0.026	0.072	0.025	1.011	0.384	0.536***	0.300***
M_PPE_GROWTH	0.038	−0.034	−0.018**	−0.067**	0.056	0.028	0.024***	0.011***
M_LEV	0.223	0.131	0.241	0.154**	0.216	0.194	0.183***	0.186***
M_MB	1.814	0.802	10.737	1.250**	1.579	1.304	1.707*	1.678***
M_RETURN	0.242	−0.021	0.242	0.011	0.114	0.049	0.112	0.062
M_RET_VOL	0.223	0.175	0.223	0.180	0.133	0.116	0.130	0.116
M_SHARE_ISSUE	0.752	1.000	0.836***	1.000	0.684	0.750	0.721***	0.750
M_LOSS	0.550	0.592	0.625***	0.667**	0.308	0.250	0.343***	0.333***
M_LITIGATE	0.342	0.142	0.376	0.334***	0.261	0.200	0.286**	0.259***
M_BANKRUPTCY	−2.268	1.010	−3.741***	−0.578***	1.742	1.946	1.754	1.837***
M_#_OPER_SEGS	1.072	1.000	1.048	1.000	1.331	1.000	1.283	1.072***
M_#_GEO_SEGS	1.810	1.000	1.869	1.583***	2.343	2.000	2.452**	2.333***

(continued on next page)

TABLE 1 (continued)

Panel E: Office-Level Audit Failures across SEC Regional Office Locations

		AUD_FAIL_X = 0	*AUD_FAIL_0* = 1		*AUD_FAIL_10* = 1	
SEC Region	**SEC Regional Office**	**Number**	**Number**	**% of Total In Regional Office**	**Number**	**% of Total In Regional Office**
Northeast	New York	156	40	20.4%	18	10.3%
	Philadelphia	583	189	24.5	127	17.9
	Boston	248	74	23.0	47	15.9
Midwest	Chicago	621	218	26.0	122	16.4
Southeast	Atlanta	517	139	21.2	72	12.2
	Miami	96	22	18.6	8	7.7
Central	Dallas-Fort Worth	139	53	27.6	32	18.7
	Denver	282	122	30.2	81	22.3
	Salt Lake City	53	32	37.6	19	26.4
Pacific	San Francisco	498	259	34.2	164	24.8
	Los Angeles	90	41	31.3	24	21.1
	Total	3,283	1,189	26.6%	714	17.9%

*, **, *** Indicate significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

Big 4 indicates that the auditor office is that of a Big 4 audit firm in year *t*. Non-Big 4 indicates that the auditor office is that of a non-Big 4 audit firm in year *t*. The auditor office-year test variable *AUD_FAIL_X* is coded 1 when at least one low-quality audit failure occurs within the same office of a company's external auditor during year *t*, and 0 otherwise. A low-quality audit is defined as existing when a client company restates net income downward by a material amount subsequent to the audit. *X* refers to the materiality level of the restatement (i.e., 0 for a greater than 0 percent downward restatement of net income and 10 for a greater than 10 percent downward restatement of net income). "*M_*" before a variable indicates that it is the median value of the variable across all audit clients in an auditor office during a year.

See Appendix A for all variable definitions.

most recent, restated net income number. We then scale this by the absolute value of the originally released net income number to obtain a restatement percentage, either positive or negative.[7] Finally, we cross-map the downward restatements from Compustat to the Audit Analytics database and use only those restatements that are in both databases.

We initially define an audit failure as having occurred when a client company restates net income *downward* by any amount compared to the originally reported value.[8] We also examine a second and more extreme restatement threshold of 10 percent (or more) of originally reported net income to ensure the results are robust to a higher materiality level. It turns out the results are consistent across both thresholds. Downward restatements indicate that the company's originally released net income was "too high" as originally audited. Given that auditors are most concerned with overstatements of net income due to liability concerns (Basu 1997; Kothari et al. 1989; Skinner 1994), and given that income-increasing accruals are more likely to result in auditor reporting conservatism (Francis and Krishnan 1999), we consider only an overstatement of originally reported net income to be an audit failure. While we consider only downward restatements to be an audit failure in our main analyses, all results are similar if we re-code an office with either downward or upward restatements of net income as being an audit failure.

The year of an audit failure is the fiscal year in which the client's restated net income was *originally* reported in the 10-K. This identification is important because we are testing whether the existence of an audit failure in a given year for an auditor office is indicative of other lower quality audits throughout that office in the same fiscal year (i.e., a contagion effect). Auditor office locations that have one or more clients with a downward restatement in net income for a particular year are coded 1 for the test variable AUD_FAIL_X, where X indicates the particular percentage restatement threshold being used in a particular model: i.e., $X = 0$ for any downward restatement of net income (0 percent threshold), and $X = 10$ for 10 percent or greater downward restatements of net income. The control sample of auditor office-years with no downward restatements of net income in a particular year are coded 0 ($AUD\_FAIL\_X = 0$).

Table 1, Panel C summarizes the coding of office-year audit failures. For Big 4 accounting firms at the 0 percent materiality threshold, 876 of 2,475 total office-years are coded 1, indicating the presence of one or more downward restatements of clients' earnings in 35.4 percent of office-years. For non-Big 4 firms at the 0 percent materiality threshold, 313 of 1,997 total office-years (15.7 percent) are coded 1. Untabulated results show that of the 313 non-Big 4 office-years with at least one audit failure, 217 of these office-years have exactly one audit failure and 96 office-years (31 percent) have more than one failure: 59 offices have two failures, 27 have three failures, eight have four failures, and two have five failures. For the 876 Big 4 office-years with audit failures, 414 have a single failure and 462 (52 percent) have more than one failure: 237 offices have two failures, 103 offices have three failures, 49 offices have four failures, 34 offices have five failures, and 39 have more than five failures. For Big 4 offices in particular, the presence of one audit failure indicates a strong likelihood of others. In total, 30 office-years have more than five audit failures, and the maximum is 17 failures in a single office.

---

[7] Scaling by the absolute value of the originally released net income value ensures that all decreases (increases) in net income due to the restatement are calculated to be a negative (positive) percentage restatement.

[8] A restatement may have occurred for any of the company's four fiscal quarters for a fiscal year. Quarter-end financial statements for the first three quarters are typically reviewed instead of audited. However, given that fiscal year-end financial statements are always audited, and given that fiscal year-end net income includes cumulative net income for all four fiscal quarters, each quarter is, in effect, audited at year-end. Therefore, a restatement of net income for any of the company's fiscal quarters can be considered a restatement of *audited* annual earnings.

Table 1, Panel D presents office-level characteristics and compares offices with and without audit failures. All variables, except *OFFICE_SIZE*, are calculated by taking the mean and median value across all client firms within an office each year. Panel C shows statistically significant differences between offices with/without audit failures for about two-thirds (three quarters) of the variables for non-Big 4 (Big 4) offices, indicating the need to control for these client characteristics in the regression models. Finally, Panel E reports the number of auditor offices with and without audit failures located in each of the SEC's 11 regional offices/districts. Kedia and Rajgopal (2011) find that geography is important in explaining corporate misreporting and SEC enforcement activity. A Kolmogorov-Smirnov test indicates that the distribution of auditor offices with failures versus offices without audit failures is significantly different ($p < 0.01$) across the 11 SEC regions. In particular, there is a higher rate of auditor office failures in the Central and Pacific regions (Dallas-Fort Worth, Denver, Salt Lake City, San Francisco, and Los Angeles) as all of these regions are above the sample mean. To control for potential regional differences, we include SEC regional office fixed effects in our models, as described later.

Empirical Model

The OLS regression model in Equation (1) is estimated for separate samples of Big 4 and non-Big 4 client companies to test if a contagion effect exists in auditor offices:

$$\begin{aligned} \mathit{ABS\_ABN\_ACC} \text{ or } \mathit{ABN\_ACC} = {} & \beta_0 + \beta_1 \mathit{AUD\_FAIL\_X} + \beta_2 \mathit{OFFICE\_SIZE} + \beta_3 \mathit{RISK\_PORT} \\ & + \beta_4 \mathit{CITY\_IND\_EXP} + \beta_5 \mathit{NAT\_IND\_EXP} + \beta_6 \mathit{INFLUENCE} \\ & + \beta_7 \mathit{SIZE} + \beta_8 \mathit{LAG\_TOT\_ACC} + \beta_9 \mathit{CFO} + \beta_{10} \mathit{CFO\_VOL} \\ & + \beta_{11} \mathit{SALES\_GROWTH} + \beta_{12} \mathit{SALES\_VOL} \\ & + \beta_{13} \mathit{PPE\_GROWTH} + \beta_{14} \mathit{LEV} + \beta_{15} \mathit{MB} + \beta_{16} \mathit{RETURN} \\ & + \beta_{17} \mathit{RET\_VOL} + \beta_{18} \mathit{SHARE\_ISSUE} + \beta_{19} \mathit{LOSS} \\ & + \beta_{20} \mathit{LITIGATE} + \beta_{21} \mathit{BANKRUPTCY} + \beta_{22} \mathit{\#\_OPER\_SEGS} \\ & + \beta_{23} \mathit{\#\_GEO\_SEGS} + \mathit{Year\ Fixed\ Effects} \\ & + \mathit{Industry\ Fixed\ Effects} + \mathit{SEC\ Regional\ Office\ Fixed\ Effects} \\ & + \varepsilon, \end{aligned} \tag{1}$$

where the dependent variable *ABS_ABN_ACC* (*ABN_ACC*) is the absolute value (signed value) of a company's abnormal accruals in year *t*, controlling for concurrent performance using a modified Jones model (Dechow et al. 1995; Jones 1991; Kothari et al. 2005). We analyze both absolute abnormal accruals and the subsample of income-increasing abnormal accruals because Hribar and Nichols (2007) demonstrate that the analysis of absolute accruals may be problematic due to a correlated omitted variable problem. The calculation of abnormal accruals is detailed in the next subsection.

Other auditor office characteristics are controlled for as prior research shows these to be important. Office size is controlled for because Francis and Yu (2009) and Choi et al. (2010) show that Big 4 office size is negatively associated with client abnormal accruals. Consistent with their studies, the variable *OFFICE_SIZE* is the natural log of the total dollar amount of audit fees charged to all audit clients within an auditor office in year *t*. A dichotomous version of this variable is also used as a test variable in Table 5 to investigate whether auditor office size affects the extent to which an audit failure is indicative of a contagion effect (test of H2).

We also control for the average clientele portfolio risk within an auditor office (*RISK_PORT*) to mitigate the concern that client-specific characteristics may be driving either the likelihood that at least one restatement occurs within an office, or the level of abnormal accruals of those clients, or both. We compute an office's client portfolio risk by first calculating the median level of client assets, leverage,

and return on assets within each office-year, similar to prior research (Johnstone and Bedard 2004).[9] We then standardize these values, which results in a mean of 0 and standard deviation of 1 (for each variable) so as to not under-/over-weight any individual variable. Finally, we add together the standardized mean values of assets and return on assets, subtract the mean value of leverage, and then multiply this sum by −1 (so that a higher value reflects a riskier portfolio) to obtain the final value of *RISK_PORT*. Results are similar if mean values are used, and we make no prediction for the sign on this variable.

We create an additional office-level test variable for H3, which examines if the percentage of audits conducted within a Big 4 office where that office is the city-level industry leader, has an impact on the contagion effect. The variable *OFFICE_EXP_#* is measured as the number of audits conducted in a Big 4 office in a year where that office is the city-level industry leader, scaled by the total number of audits conducted by the office in the year.

Engagement-specific auditor industry expertise is controlled using both city- and national-level measures in which the industry leader (auditor with the largest dollar amount of audit fees) is considered to be the industry expert (Francis et al. 2005; Reichelt and Wang 2010). Only Big 4 auditors are city and/or national industry leaders. National (city) industry leadership is based on each audit firm's market share of audit fees in a two-digit SIC category in the United States (within a two-digit SIC category in a specific city). Following Francis et al. (2005) and Reichelt and Wang (2010), we define a city using the Metropolitan Statistical Area (MSA) as classified by the U.S. Census Bureau. Auditor cities are collected from Audit Analytics and are then categorized by MSA using the U.S. Census Bureau's MSA cross-map.[10] Both city and national industry leaderships are recalculated each year. The variable *CITY_IND_EXP* is coded 1 if the auditor on a specific client engagement is the city-specific market share leader in terms of audit fees in a given year, and *NAT_IND_EXP* is coded 1 if the auditor is the national market share leader in terms of audit fees in a given year. Results on the association between abnormal accruals and both city and national industry leadership are mixed in prior research so we do not make a prediction for these variables (Francis and Yu 2009; Reichelt and Wang 2010).

The variable *INFLUENCE* is the total dollar amount of audit and nonaudit fees charged to a specific client in year *t*, scaled by the total fees charged by the auditor office in a year. Francis and Yu (2009) include this variable to control for the possibility that a specific client that provides a relatively high percentage of total fees to an auditor office may affect auditor objectivity and audit quality for that client. In most of their analyses this variable is not significant, so we do not predict a sign for the coefficient on *INFLUENCE*.

Firm-level variables used in prior studies are included in all analyses to control for the various characteristics that affect a company's level of abnormal accruals (see Appendix A for detailed definitions for all control variables). Based on prior research, we expect *SIZE*, *LAG_TOT_ACC*, *CFO*, *LOSS*, and *BANKRUPTCY* to be negatively associated with abnormal accruals, while we expect *CFO_VOL*, *SALES_GROWTH*, *SALES_VOL*, *PPE_GROWTH*, *MB*, *RET_VOL*, and *LITIGATE* to be positively associated with abnormal accruals (Choi et al. 2010; Francis and Yu 2009; Hribar and Nichols 2007; Reichelt and Wang 2010). We do not predict a sign for *LEV*, *RETURN*, *SHARE_ISSUE*, *#_OPER_SEGS*, and *#_GEO_SEGS* due to absent or conflicting results

[9] Johnstone and Bedard (2004) include additional variables in calculating their risk portfolio measure. However, these variables were obtained through a questionnaire specific to their study. We use variables that are available in Compustat. Further, Johnstone and Bedard (2004) do not include client assets as a risk variable, although they do include it as a control variable in their analyses.

[10] The MSA cross-map (2008 definition) is available at: http://www.census.gov/population/www/metroareas/metroarea.html. For cities not listed on the cross-map, we hand-collect the closest MSA using the 2008 map available at the website listed above and Google Maps. We thank Brett Kawada and Sarah Stein for their help in this hand-collection.

in prior studies (Francis and Yu 2009).[11] As in prior research, we include year and industry fixed effects. In addition, SEC regional office fixed effects are used to control for geographic patterns in misreporting and detection by the SEC (Kedia and Rajgopal 2011).

Abnormal Accruals

Firm-year abnormal accruals are calculated using a modified Jones model (Dechow et al. 1995; Jones 1991), controlling for concurrent performance (Kothari et al. 2005) within industry-year groups for separate samples of Big 4 and non-Big 4 clients, where industries are defined by a company's two-digit SIC code. The model in Equation (2) is estimated separately for each industry-year-auditor group, and requires a minimum of 20 observations:[12]

$$TOT\_ACC = \alpha_0 + \alpha_1(1/ASSETS) + \alpha_2(\Delta SALES - \Delta AR) + \alpha_3 PPE + \alpha_4 ROA + \varepsilon. \quad (2)$$

The variable *TOT_ACC* is calculated as a company's net income before extraordinary items less cash flows from operations; *ASSETS* is a company's total assets at the end of year $t-1$; *SALES* is a company's sales in year t and $t-1$ scaled by lagged total assets; *AR* is a company's net total receivables at the end of year t and $t-1$ scaled by lagged total assets; *PPE* is net property, plant, and equipment at the end of year t scaled by lagged total assets; and *ROA* is net income in year t scaled by lagged total assets. Equation (2) is estimated separately for clients of Big 4 and non-Big 4 accounting firms (for separate industry-year subsamples) as these clients exhibit different operating and accounting characteristics (Francis et al. 1999a), although the results are qualitatively the same if Equation (2) is estimated for the full sample.

Equation (2) uses only those firm-year observations that do not have an earnings restatement, either income-increasing or income-decreasing, as the inclusion of companies with misstated earnings could bias the calculation of the coefficient parameters. We then apply these parameter values to firm-year observations in the treatment and control samples to derive expected accruals. Abnormal accruals are the difference between expected and actual accruals.

Descriptive Statistics

Table 2, Panel A presents the descriptive statistics of the variables in the study. *ABS_ABN_ACC* has a mean (median) value of 0.091 (0.052), which is similar to other studies (Reichelt and Wang 2010; Reynolds and Francis 2000). The mean value of *ABN_ACC* for all observations is 0 by construction (Kothari et al. 2005), and the median value is also close to 0 (–0.003). The office-level test variable *AUD_FAIL_0* is coded 1 when at least one client restates net income downward by any amount within an auditor office in a year, and 0 otherwise. There are 4,472 total office-years in the sample, and the mean value of 0.265 indicates that 26.5 percent of auditor office-years have an audit failure at the 0 percent materiality threshold). At the 10 percent threshold (*AUD_FAIL_10*), 16.0 percent of auditor office-years have an audit failure.

---

[11] Francis and Yu (2009) also include the variable *TENURE* that indicates whether a company has been audited by the same audit firm for at least three years, based on Johnson et al. (2002). They are able to include this variable for their entire sample because they begin their analysis in the year 2003. However, given that this variable requires two years of lagged data to compute, its inclusion would force us to eliminate observations in the years 2000 and 2001 because specific auditor information is not available in Audit Analytics prior to 2000. Given the effect on sample size, we do not present analyses including *TENURE*. However, we note that all results are qualitatively the same when *TENURE* is included for a reduced sample.

[12] Observations for which any value of the variables in Equation (2) is above the 0.99 value or below the 0.01 value of all companies are excluded from the calculation of parameter values for Equation (2) to mitigate the effect of these extreme values on the calculation of expected accruals. However, these companies are included in the final sample of 22,626 company-year observations.

TABLE 2

Descriptive Statistics

Panel A: Distributional Properties of Variables

Variable	n	Mean	Std. Dev.	25%	Median	75%
ABS_ABN_ACC	22,626	0.091	0.115	0.023	0.052	0.108
ABN_ACC	22,626	0	0.126	−0.056	−0.003	0.049
AUD_FAIL_0	4,472	0.265	0.440	0	0	1.000
AUD_FAIL_10	4,472	0.160	0.366	0	0	0
B4	22,626	0.803	0.397	1.000	1.000	1.000
OFFICE_SIZE (n = office-years)	4,472	7,510	18,200	201	1,136	5,664
RISK_PORT (n = office-years)	4,472	0.669	2.725	−1.012	−0.032	1.487
OFFICE_EXP_% (Big 4 only)	2,475	0.682	0.329	0.400	0.740	1.000
CITY__IND_EXP (Big 4 only)	18,164	0.505	0.499	0	1.000	1.000
NAT_IND_EXP	22,626	0.185	0.389	0	0	0
INFLUENCE	22,626	0.243	0.495	0.020	0.066	0.234
SIZE	22,626	1,457	3,632	43	202	927
LAG_TOT_ACC	22,626	−0.151	1.471	−0.115	−0.057	−0.015
CFO	22,626	−0.052	0.488	−0.100	0.056	0.151
CFO_VOL	22,626	0.165	0.232	0.045	0.089	0.182
SALES_GROWTH	22,626	0.135	0.384	−0.045	0.075	0.219
SALES_VOL	22,626	1.108	3.047	0.041	0.182	0.781
PPE_GROWTH	22,626	0.077	0.386	−0.112	0.007	0.165
LEV	22,626	0.207	0.224	0.005	0.146	0.328
MB	22,626	2.197	3.224	0.126	1.349	2.693
RETURN	22,626	0.236	0.989	−0.283	0.031	0.407
RET_VOL	22,626	0.165	0.140	0.085	0.128	0.199
SHARE_ISSUE	22,626	0.733	0.441	0	1.000	1.000
LOSS	22,626	0.403	0.490	0	0	1.000
LITIGATE	22,626	0.315	0.464	0	0	1.000
BANKRUPTCY	22,626	0.399	5.740	0.508	1.791	2.819
#_OPER_SEGS	22,626	1.241	1.046	1.000	1.000	1.000
#_GEO_SEGS	22,626	2.486	2.087	1.000	2.000	3.000

Panel B: Differences in Means/Medians of Abnormal Accruals in Non-Big 4 Auditor Offices

	ABS_ABN_ACC			*ABN_ACC* > 0		
	Mean	**Median**	**n**	**Mean**	**Median**	**n**
AUD_FAIL_X = 0	0.139	0.079	3,193	0.122	0.080	1,642
AUD_FAIL_0 = 1	0.158***	0.087***	1,269	0.132*	0.086	602
AUD_FAIL_10 = 1	0.159**	0.088*	853	0.128	0.086	398

(continued on next page)

Approximately 80 percent of companies in the sample use a Big 4 auditor, which is consistent with prior research (Francis et al. 1999a). *OFFICE_SIZE* is presented in Table 2 in raw form as is the total dollar amount of audit fees (in \$ thousands) charged by an auditor office-year. The mean (median) value of audit fees charged is about \$7.5 million (\$1.1 million). The mean (median) values of the variable that measures an office's risk portfolio (*RISK_PORT*) are 0.669 (−0.032). The variable *OFFICE_EXP_%* is the percentage of total audits conducted by an office in a year where the Big 4

TABLE 2 (continued)

Panel C: Differences in Means/Medians of Abnormal Accruals in Big 4 Auditor Offices

	ABS_ABN_ACC			*ABN_ACC* > 0		
	Mean	**Median**	**n**	**Mean**	**Median**	**n**
AUD_FAIL_X = 0	0.072	0.045	6,846	0.071	0.044	3,344
AUD_FAIL_0 = 1	0.081***	0.050***	11,318	0.082***	0.049***	5,324
AUD_FAIL_10 = 1	0.084***	0.052***	7,839	0.086***	0.051***	3,701

*, **, *** Indicate significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.
ABS_ABN_ACC is the absolute value of a company's abnormal accruals as calculated in Kothari et al. (2005). *ABN_ACC* is the signed value of a company's abnormal accruals as calculated in Kothari et al. (2005). *AUD_FAIL_X* is 1 when at least one low-quality audit occurs within the same office of a company's external auditor during year *t*, and 0 otherwise. A low-quality audit is defined as existing when a client company restates net income downward by a material amount subsequent to the audit. *X* refers to the materiality level of the restatement (i.e., 0 for a greater than 0 percent downward restatement of net income and 10 for a greater than 10 percent downward restatement of net income). *B4* is 1 if the company hires a Big 4 auditor in year *t*, and 0 otherwise.
See Appendix A for definitions of all other variables.

auditor is the city-specific industry leader. Therefore, *OFFICE_EXP_%* is a continuous variable that is specific to each auditor-office-year observation. The mean (median) values for *OFFICE_EXP_%* over the 2,475 Big 4 auditor-office-year observations is 0.682 (0.740). This indicates that the average Big 4 office is the city-level industry leader on approximately 70 percent of its audit engagements.

The variable *CITY_IND_EXP* is a firm-year specific variable that takes on a value of 1 when a company is audited by the city-level industry leader in a year, and 0 otherwise. The mean value of *CITY_IND_EXP* is 0.505, indicating that about half of all Big 4 audit engagements in the sample are conducted by a city-level industry leader. The mean value of *NAT_IND_EXP* is 0.185 indicating that for about 18 percent of Big 4 audits, the auditor is classified as the national industry leader.

The variable *INFLUENCE* has a mean (median) value of 0.243 (0.066) indicating that the average client company represents 24.3 percent of the total fees charged to all clients of an office. The median value is only 6.6 percent, which indicates that this variable is skewed and that some particularly highly influential clients are driving the mean.[13]

Table 2, Panels B and C present univariate results comparing the mean and median values of companies' absolute and income-increasing abnormal accruals in offices with at least one audit failure compared to offices with zero audit failures. Panel B presents t-tests (rank-sum tests) for differences in means (medians) for non-Big 4 offices. The values for *AUD_FAIL_X* = 0 are the mean and median levels of client company absolute abnormal accruals (*ABS_ABN_ACC*) and income-increasing abnormal accruals (*ABN_ACC* > 0) audited by auditor offices with no audit failures in a year. *AUD_FAIL_0* = 1 and *AUD_FAIL_10* = 1 present the same values for office-years with at least one audit failure at the 0 and 10 percent threshold levels, respectively. Results at both thresholds indicate the mean and median level of client absolute abnormal accruals are significantly larger in non-Big 4 offices with at least one audit failure compared to non-Big 4 offices with no audit failures: at the 10 percent threshold, mean abnormal accruals are larger by 2.0 percent of lagged assets. Results for income-increasing abnormal accruals are less clear as only one out of four differences is significant. However, this is before controlling for other important factors that likely affect abnormal accruals. Panel C presents mean/median values for Big 4 offices and all differences are significant ($p < 0.01$). At the 10

[13] All results on our analyses of a contagion effect are virtually identical if we use either the log or the square root of *INFLUENCE* as a control variable.

percent threshold, mean abnormal accruals are larger by 1.2 percent of lagged assets, and mean income-increasing accruals are larger by 1.6 percent of lagged assets. Taken together, the results in both panels present univariate evidence consistent with a contagion effect in non-Big 4 and Big 4 offices.

Untabulated Pearson and Spearman correlations show the correlations among the independent variables are all below 0.50, indicating that multicollinearity is not likely to be of concern. This conclusion is supported by variance inflation factors in the model estimations that are all less than 6.5, well below the threshold of 10 suggested in Kennedy (1992).

IV. RESULTS

Before testing the study's three hypotheses, we first examine if an audit failure in a prior office-year increases the likelihood of observing an audit failure (misstated client earnings) in a subsequent office-year. Given that auditor office-years are the unit of analysis, to control for office-specific clientele characteristics we calculate the median level of all firm-level control variables used in Equation (1) for clients of the office-year.[14] We do not predict a sign on the office-level control variables as we know of no prior research to rely on that predicts client restatements at the office level. However, it is important to control for office-level characteristics that may affect the likelihood that one or more clients of an office issues a restatement. Table 3 reports this analysis at the 0 percent materiality threshold using a Probit regression.[15] For both Big 4 and non-Big 4 auditors, an audit failure in office-year *t* is more likely to occur if there were prior failures in years *t*–1 through *t*–5. Thus, when we observe an audit failure in an office, there is a significant likelihood there will be future (new) audit failures for up to five subsequent years. Results are similar at the 10 percent materiality threshold. It appears that certain engagements offices experience "serial problems" with audit quality.[16]

This result may seem surprising because a restatement is publicly disclosed and one might expect an audit firm to respond quickly to such an event by strengthening quality-control procedures within the office.[17] Even though it takes, on average, two years for a restatement to occur, it is remarkable that even after five years there is still a significant likelihood of another audit failure in the office. The lack of timely remediation could be due to the perception that a specific restatement was due to unique client-specific characteristics of the audit engagement (poor internal controls, weak financial reporting oversight, or even fraud that was not detected), or to idiosyncratic engagement-specific auditor characteristics, rather than suggestive of a more systemic problem in office-level audit quality. Our results suggest the possibility that audit firms can use client restatements as another input to their national quality-control procedures to identify problems and to improve audit quality in offices throughout the firm.

---

[14] Results are very similar if mean values are used in place of median value for all office-level control variables.

[15] For this analysis we delete consecutive multi-year restatements by the same company as such restatements would bias in favor of the test. For a specific company we keep only the first restatement and recode *AUD_FAIL_X* accordingly for the analysis in Table 3.

[16] For both Tables 3 and 4 we also run models that pool Big 4 and non-Big 4 clients together. We run these pooled models in two ways: (1) We add a *B4* indicator variable and the interaction between *B4* and all five lagged versions of *AUD_FAIL_0* (Table 3), and a *B4* indicator variable along with the interaction between *B4* and *AUD_FAIL_X* (Table 4) and (2) we add the *B4* indicator variable as well as its interaction with all other control variables (except fixed effects) in Tables 3 and 4. In all cases inferences drawn from these results are virtually identical to those reported in the study. Further, the negative and significant coefficients on *B4* in these estimations indicate that Big 4 offices are less likely to have an audit failure overall (Table 3), and clients of Big 4 offices exhibit lower levels of abnormal accruals overall (Table 4).

[17] Correlations among the office-level variables show that fewer (more) audit failures occur, on average, in larger Big 4 (non-Big 4) offices, and in offices that have a smaller percentage of audits in industries for which the office is the city-level industry leader. The average clientele risk of an office (*RISK_PORT*) does not seem to be related to audit failures after controlling for specific client risk characteristics in our multivariate analyses. However, the univariate correlation indicates that the average clientele risk of an office (*RISK_PORT*) is negatively correlated with the presence of at least one audit failure. This suggests that auditors adjust their audit procedures to be more stringent when clients are higher risk, which is consistent with Generally Accepted Auditing Standards.

TABLE 3

Current Audit Failures in an Office Predicted by Past Audit Failures
Auditor Office-Level Analysis
(dependent variable = *AUD_FAIL_0*)

Variable	Pred.	Non-B4	B4	Non-B4	B4	Non-B4	B4	Non-B4	B4	Non-B4	B4
Test Variables											
LAG_1_AUD_FAIL_0	+	0.103***	0.286***								
LAG_2_AUD_FAIL_0	?			0.117***	0.194***						
LAG_3_AUD_FAIL_0	?					0.046*	0.183***				
LAG_4_AUD_FAIL_0	?							0.093***	0.146***		
LAG_5_AUD_FAIL_0	?									0.127***	0.120***
(p-value)		(0.000)	(0.000)	(0.000)	(0.000)	(0.072)	(0.000)	(0.003)	(0.000)	(0.000)	(0.000)
Control Variables											
OFFICE_SIZE	+	0.002***	0.003***	0.002***	0.004***	0.002***	0.004***	0.002***	0.004***	0.002***	0.004***
M_CITY_IND_EXP	?		0.033		0.008		0.012		−0.008		−0.017
M_NAT_IND_EXP	?		0.018		0.031		0.013		0.009		−0.011
M_INFLUENCE	?	−0.218***	−0.590***	−0.255***	−0.613***	−0.292***	−0.571***	−0.287***	−0.553***	−0.277***	−0.534***
M_AB_ACC	?	0.006	−0.095	0.008	0.031	0.000	0.103	−0.017	−0.322	−0.009	−0.221
M_SIZE	?	0.005	0.030*	0.000	0.030*	−0.005	0.036*	−0.011	0.032*	−0.005	0.023*
M_LAG_TOT_ACC	?	−0.016	0.412**	−0.023	0.311*	0.018	0.215**	0.026	0.326**	0.036	−0.011
M_CFO	?	−0.005	0.036	−0.001	0.130	−0.003	0.134	0.019	0.142	0.025	0.334**
M_CFO_VOL	?	0.018	0.242	0.011	0.202	0.036	0.355	0.072*	0.518*	0.041	0.551*
M_SALES_GROWTH	?	0.018*	0.008	0.028	−0.004	0.040*	0.000	0.062*	0.136	0.062*	0.113
M_SALES_VOL	?	0.025	0.006	0.095**	0.004	0.132**	0.001	0.136**	−0.005	0.097*	−0.004
M_PPE_GROWTH	?	−0.004	−0.067	−0.007	−0.148	−0.001	−0.111	0.030	−0.274*	0.039	−0.205*
M_LEV	?	−0.003	−0.088	−0.025	−0.089	−0.028	−0.089	−0.003	−0.072	−0.013	−0.073
M_MB	?	−0.002	−0.016	−0.006*	−0.024*	−0.007*	−0.019	−0.014***	−0.018	−0.020***	−0.032*
M_ RETURN	?	0.006**	0.036	0.010	0.044	0.012	0.022	0.010	0.017	0.014	−0.008
M_RET_VOL	?	0.005	0.182	0.001	0.002	−0.040	0.204	−0.138	0.507	0.004	0.264
M_SHARE_ISSUE	?	0.049**	0.060	0.074***	0.080	0.079**	0.034	0.038	0.014	0.017	0.027
M_LOSS	?	−0.014	−0.045	0.001	−0.007	0.010	0.019	0.051	−0.007	0.072	0.016
M_LITIGATE	?	0.015	0.044	0.015	0.043	0.018	−0.012	0.003	−0.038	0.022	−0.045
M_BANKRUPTCY	?	−0.001	−0.012*	−0.001	−0.011*	−0.001	−0.007*	−0.000	0.003	0.002	0.000

(continued on next page)

TABLE 3 (continued)

Variable	Pred.	Non-B4	B4	Non-B4	B4	Non-B4	B4	Non-B4	B4	Non-B4	B4
M_#_OPER_SEGS	?	−0.009	−0.014	0.002	−0.015	−0.007	−0.022	−0.005	−0.057*	−0.011	−0.028
M_#_GEO_SEGS	?	−0.008	0.003**	−0.010	0.006**	−0.020**	0.006**	−0.027**	0.003	−0.041***	0.003
Intercept	?	0.070	0.050	0.051	0.103	0.033	0.127*	−0.002	0.124	0.101	0.106
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SEC Regional Office Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
n (treatment)		314	814	251	740	197	648	144	539	94	412
n (control)		1,451	1,345	1,233	1,065	993	858	762	689	537	555
Model p-value		< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Pseudo R^2		29.8%	33.9%	28.4%	30.9%	26.9%	29.0%	26.5%	28.1%	26.5%	27.0%

*, **, *** Indicate significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.
t-statistics are calculated based on robust standards errors clustered at the auditor-office level. *AUD_FAIL_0* is 1 when at least one audit failure (defined as a downward restatement of income) occurs in the office of the company's external auditor during year *t*, and 0 otherwise. "Lag" denotes the value of *AUD_FAIL_0* in years *t*–1, *t*–2, *t*–3, *t*-4, and *t*–5, respectively. "*M_*" before a variable indicates that it is the median (or mean for dichotomous variables) value in an office-year.
See Appendix A for definitions of control variables.

Tests of H1—Contagion in Non-Big 4 and Big 4 Auditor Offices

The results in Table 3 provide evidence of longitudinal contagion of audit failures in offices over time. H1 tests if there is cross-sectional contagion on the quality of audited earnings for other concurrent clients in office-years with an audit failure. Table 4, Panel A presents the first set of regression results testing H1 where the dependent variable is a company's level of absolute abnormal accruals (*ABS_ABN_ACC*). The p-values on the test variable of interest (*AUD_FAIL_X*) are reported conservatively as two-tailed values, even though directional predictions are being made with respect to contagion effects. All models are significant at $p < 0.01$ with R^2s around 47 percent for clients of non-Big 4 offices, and 27 percent for clients of Big 4 offices. These and all subsequent models include industry (two-digit SIC codes), year, and SEC regional offices fixed effects, and t-statistics are based on standard errors clustered at the auditor-office level.[18] Auditor and company-level control variables are generally consistent with predictions.

In Table 4, Panel A, the positive and significant ($p < 0.05$, two-tailed) coefficients on *AUD_FAIL_X* for Big 4 clients for both thresholds indicate that clients audited by Big 4 offices with at least one audit failure have higher absolute abnormal accruals, on average, compared to Big 4 offices where no audit failures occurred. These results support H1 and are consistent with the univariate tests in Table 2, Panel C, and provide evidence that a contagion effect exists in Big 4 auditor offices. In contrast, for non-Big 4 offices the coefficients are not significant at either threshold, indicating that absolute abnormal accruals are not different in non-Big 4 offices with an audit failure compared to offices with no failures.

Table 4, Panel B analyzes whether these results hold when looking only at companies that exhibit positive or income-increasing abnormal accruals. Results for Big 4 auditors are similar to Panel A, with the coefficients on *AUD_FAIL_X* being significant at the 0.01 level (two-tailed). The results for non-Big 4 offices are significant at both materiality thresholds ($p < 0.10$, two-tailed) indicating that income-increasing abnormal accruals are higher in non-Big 4 offices with at least one audit failure. Taken together, the results in both panels of Table 4 indicate that contagion occurs in the offices of both non-Big 4 and Big 4 accounting firms, with the caveat that results are not significant for absolute abnormal accruals in non-Big 4 offices. We conclude that the presence of a downward material client restatement of net income provides a method for assessing the earnings quality of clients in office-years by analyzing very simple, publicly available information.

The results in Table 4 are also economically significant. Given that *AUD_FAIL_X* is an indicator variable, the results in the fourth model of Panel A indicate that a Big 4 client company's level of absolute abnormal accruals is higher by a magnitude of 0.004 when audited by an office where at least one audit failure occurs (at the 0 percent threshold). This magnitude represents an increase of 5.1 and 4.7 (8.3 and 6.5) percent over Big 4 client companies' mean (median) values of absolute abnormal and total accruals, respectively.[19] For Panel B, the magnitude of the coefficient on *AUD_FAIL_X* in the third model of 0.009 represents a 7.2 and 4.4 (11.1 and 12.9) percent increase over non-Big 4 client companies' mean (median) values of income-increasing abnormal and total accruals, respectively.[20] Similarly, the 0.007 magnitude in the fourth model represents a 9.0 and 8.2 (14.9 and 11.4) percent increase over Big 4 client companies' mean (median) values of

---

[18] We cluster standard errors at the auditor-office level instead of the company level because our variables of interest vary at the office level, not the company level. Therefore, standard errors that are not clustered may be inflated due to including multiple observations of the same auditor office in the sample. However, if standard errors are clustered at the company level instead, all results are very similar in terms of the statistical significance on all test variables.

[19] The untabulated mean/median values of absolute abnormal and (total) accruals for Big 4 client companies are 0.0778/0.0481 and (–0.0851/–0.0616).

[20] The untabulated mean/median values of income-increasing abnormal and (total) accruals for non-Big 4 client companies are 0.1246/0.0811 and (–0.2063/–0.0700).

TABLE 4

Audit Failures within an Auditor Office and Abnormal Accruals
Client Company-Level Analysis

Panel A: Dependent Variable = *ABS_ABN_ACC*

		Audit Failure Threshold (*AUD_FAIL_X*)			
		> 0%		> 10%	
Variable	**Pred.**	**Non-B4**	**B4**	**Non-B4**	**B4**
Test Variable					
AUD_FAIL_X	+	0.004	0.004**	0.002	0.004**
(p-value)		(0.303)	(0.011)	(0.683)	(0.015)
Control Variables					
OFFICE_SIZE	–	0.000	–0.002**	–0.001	–0.003***
RISK_PORT	?	–0.001	–0.001	0.001	–0.001
CITY_IND_EXP	?		–0.000		–0.004
NAT_IND_EXP	?		0.001		–0.003
INFLUENCE	?	–0.001	0.001	–0.001	0.000
SIZE	–	–0.014***	–0.008***	–0.013***	–0.008***
LAG_TOT_ACC	–	–0.049***	–0.019**	–0.053***	–0.016*
CFO	–	–0.060***	–0.036***	–0.059***	–0.036***
CFO_VOL	+	0.089***	0.068***	0.092***	0.068***
SALES_GROWTH	+	0.016***	0.012***	0.017***	0.011***
SALES_VOL	+	0.015**	0.001***	0.013*	0.001***
PPE_GROWTH	+	0.001	0.017***	0.001	0.019***
LEV	?	–0.008	–0.019***	–0.008	–0.023***
MB	+	0.001	0.001***	0.001*	0.002***
RETURN	?	0.002*	0.002***	0.001	0.002**
RET_VOL	+	0.044***	0.043***	0.048***	0.051***
SHARE_ISSUE	?	–0.003	–0.001	–0.003	–0.001
LOSS	–	–0.028***	–0.001	–0.028***	–0.002
LITIGATE	+	–0.003	0.002	–0.002	0.003
BANKRUPTCY	–	–0.002***	–0.003***	–0.001***	–0.004***
#_OPER_SEGS	?	–0.003	0.000	–0.002	0.000
#_GEO_SEGS	?	–0.000	–0.001***	–0.000	–0.001***
Intercept	?	0.167***	0.145***	0.150***	0.140***
Year Fixed Effects		Yes	Yes	Yes	Yes
Industry Fixed Effects		Yes	Yes	Yes	Yes
SEC Regional Office Fixed Effects		Yes	Yes	Yes	Yes
n (treatment)		1,269	11,318	853	7,839
n (control)		3,193	6,846	3,193	6,846
Model p-value		< 0.001	< 0.001	< 0.001	< 0.001
R^2		46.6 %	26.9 %	46.3 %	27.9 %

(continued on next page)

TABLE 4 (continued)

Panel B: Dependent Variable = *ABN_ACC* (> 0)

Variable	Pred.	Audit Failure Threshold (*AUD_FAIL_X*) > 0% Non-B4	> 0% B4	> 10% Non-B4	> 10% B4
Test Variable					
AUD_FAIL_X	+	0.007*	0.006***	0.009*	0.007***
(p-value)		(0.098)	(0.003)	(0.081)	(0.003)
Control Variables					
OFFICE_SIZE	–	–0.000	–0.001	–0.001	–0.001
RISK_PORT	?	0.001	0.001	0.001	0.001
CITY_IND_EXP	?		–0.000		0.001
NAT_IND_EXP	?		–0.000		–0.001
INFLUENCE	?	0.002	0.005**	0.002	0.005**
SIZE	–	–0.009***	–0.011***	–0.008	–0.012***
LAG_TOT_ACC	–	0.012**	–0.021***	0.017**	–0.016**
CFO	–	–0.036***	–0.025***	–0.033***	–0.025***
CFO_VOL	+	0.038***	0.048***	0.037***	0.048***
SALES_GROWTH	+	0.007*	0.014***	0.008**	0.013***
SALES_VOL	+	0.006	0.002***	0.004	0.002***
PPE_GROWTH	+	–0.004	0.014***	–0.001	0.014***
LEV	?	–0.007	–0.004	–0.011	–0.004
MB	+	0.001	0.001	0.001**	0.001
RETURN	?	0.005***	0.001	0.004**	0.002
RET_VOL	+	–0.001	0.059***	–0.006	0.063***
SHARE_ISSUE	?	–0.001	0.003*	–0.002	0.002
LOSS	–	–0.011***	–0.012***	–0.011***	–0.012***
LITIGATE	+	0.001	–0.003	0.001	–0.001
BANKRUPTCY	–	–0.003***	–0.003***	–0.003***	–0.004***
#_OPER_SEGS	?	–0.004	0.001	–0.004	0.001*
#_GEO_SEGS	?	0.000	–0.001**	0.000	–0.001*
Intercept	?	0.127	0.152***	0.129	0.152***
Year Fixed Effects		Yes	Yes	Yes	Yes
Industry Fixed Effects		Yes	Yes	Yes	Yes
SEC Regional Office Fixed Effects		Yes	Yes	Yes	Yes
n (treatment)		602	5,324	398	3,701
n (control)		1,642	3,344	1,642	3,344
Model p-value		< 0.001	< 0.001	< 0.001	< 0.001
R^2		29.1%	31.1%	30.3%	31.7%

*, **, *** Indicate significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.
t-statistics are calculated based on robust standards errors clustered at the auditor-office level. *ABS_ABN_ACC* is the absolute value of a company's abnormal accruals as calculated in Kothari et al. (2005). *ABN_ACC* is the signed value of a company's abnormal accruals as calculated in Kothari et al. (2005). *AUD_FAIL_X* is 1 when at least one low-quality audit failure occurs within the same office of a company's external auditor during year *t*, and 0 otherwise. A low-quality audit is defined as existing when a client company restates net income downward by a material amount subsequent to the audit. *X* refers to the materiality level of the restatement (i.e., 0 for a greater than 0 percent downward restatement of net income and 10 for a greater than 10 percent downward restatement of net income).
See Appendix A for definitions of control variables.

income-increasing abnormal and total accruals.[21] We conclude that the accrual differences in treatment and control offices are both economically and statistically significant.

In untabulated results we also examine the effect of lagged values of *AUD_FAIL_0* on current (year *t*) abnormal accruals in order to test if the contagion effect on earnings quality persists over time for auditor offices having an audit failure. These results show a significant lagged contagion effect for Big 4 auditors in years $t-1$, $t-3$, and $t-4$, and a significant lagged contagion effect for non-Big 4 auditors in years $t-2$, $t-3$, and $t-5$. Given that, for both auditor types, we find a significant contagion effect in three of the five lagged years, we conclude that there exists a greater likelihood of lower earnings quality in the current year *t* for those auditor offices with recent past audit failures. These results complement the analysis in Table 3 and further indicate that some offices have a serial problem with low-quality audits.

Tests of H2—Contagion Effect for Larger versus Smaller Offices

Table 5 investigates whether the contagion effect in auditor offices is conditioned by the size of the office (H2), and is motivated by the findings that larger Big 4 offices produce higher-quality audits. For completeness we analyze both Big 4 and non-Big 4 offices in separate tests. For this analysis we create an indicator variable *LARGE_OFFICE*, and this variable is coded 1 for auditor offices that are in the largest quartile of offices (*OFFICE_SIZE*), and 0 otherwise. For Big 4 offices, the cutoff is \$13.5 million of audit fees (using the size distribution of Big 4 offices), and the cutoff is \$667,000 in audit fees for non-Big 4 offices (using the size distribution of non-Big 4 offices). By comparing the largest quartile of auditor office size to the other 75 percent of offices within each group, the test is biased against finding a contagion effect in smaller offices because many of the offices coded as "small" are still quite large and well above median office size.

For brevity, Table 5 reports results only for the 10 percent threshold; however, signs and significance levels are comparable at the 0 percent materiality threshold. Further, we do not report the coefficients on the control variables, although all controls that are in the models in Table 4 are included. In Table 5 the test variable *AUD_FAIL_10*, by itself, determines if a contagion effect exists in the smallest 75 percent of auditor offices. The interaction coefficient *AUD_FAIL_10* ∗ *LARGE_OFFICE* tests the incremental difference for the largest quartile of offices relative to the smallest 75 percent of offices. The total contagion effect for larger offices is the sum of the coefficients on *AUD_FAIL_10* and *LARGE_OFFICE* ∗ *AUD_FAIL_10*, and an F-statistic test if the sum is different from 0 (test of H2).

In the test of absolute abnormal accruals (first two models), the coefficients on *AUD_FAIL_10* are positive and significant at $p < 0.01$ for both auditor groups, and the coefficients on the interaction term *AUD_FAIL_10* ∗ *LARGE_OFFICE* are negative and significant ($p < 0.01$ and $p < 0.05$, two-tailed, respectively) for both Big 4 and non-Big 4 offices. This means that earnings quality is lower (larger absolute abnormal accruals) for offices that are in the smallest 75 percent of the distribution for each auditor type. In contrast, there is no evidence of contagion for the largest quartile of office size for either Big 4 or non-Big 4 offices. The interaction terms are negative and significant, indicating smaller abnormal accruals relative to smaller office. The F-test of the sum of the coefficients (*AUD_FAIL_10* + *LARGE_OFFICE* ∗ *AUD_FAIL_10*) is statistically insignificant at the 0.10 level in both cases, which means the sum of the coefficients is not different from 0. We conclude that office size appears to mitigate contagion effects.

For the test of income-increasing abnormal accruals (third and fourth models), the results are comparable to those for the full sample of absolute abnormal accruals. The clients of smaller offices

[21] The untabulated mean/median values of income-increasing abnormal accruals for Big 4 client companies are 0.0780/0.0469. See footnote 23 for mean and median values of total accruals for Big 4 client companies.

TABLE 5

Audit Failures within an Auditor Office and Abnormal Accruals Dependent on Auditor Office Size

		Dependent Variable is			
		ABS_ABN_ACC		*ABN_ACC* > 0	
Variable	**Pred.**	**Non-B4**	**B4**	**Non-B4**	**B4**
Test Variables					
LARGE_OFFICE	?	0.007	0.001	−0.002	0.006**
AUD_FAIL_10	+	0.012***	0.008***	0.013**	0.011***
(p-value)		(0.005)	(0.002)	(0.024)	(0.001)
*AUD_FAIL_10 * LARGE_OFFICE*	−	−0.014***	−0.008**	−0.013	−0.010**
(p-value)		(0.001)	(0.036)	(0.196)	(0.018)
F-test [*AUD_FAIL_10* + *AUD_FAIL_10 * LARGE_OFFICE*]		1.7	0.1	0.0	0.3
(p-value)		(0.121)	(0.891)	(0.982)	(0.599)
All Control Variables		Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes
Industry Fixed Effects		Yes	Yes	Yes	Yes
SEC Regional Office Fixed Effects		Yes	Yes	Yes	Yes
n		4,046	14,685	2,040	7,045
R^2		45.1%	27.9%	29.3%	31.7%

*, **, *** Indicate significance at the 0.10, 0.05 and 0.01 levels, respectively, using two-tailed tests.
t-statistics are calculated based on robust standards errors clustered at the auditor-office level. *ABS_ABN_ACC* is the absolute value of a company's abnormal accruals as calculated per Kothari et al. (2005). *ABN_ACC* is the signed value of a company's abnormal accruals as calculated per Kothari et al. (2005). Only observations with values of *ABN_ACC* greater than 0 (income-increasing abnormal accruals) are used in the third and fourth models. *AUD_FAIL_X* is 1 when at least one audit failure occurs within the same office of a company's external auditor during year *t*, and 0 otherwise. An audit failure is defined as existing when at least one of the auditor's client companies within an auditor office subsequently restates net income downward in the future by threshold level *X*. *X* refers to the threshold level of the restatement (i.e., 0 (10) for a greater than 0 (10) percent restatement of net income). *LARGE_OFFICE* is 1 for Big 4 auditor offices that are larger than the 75th percentile value of *OFFICE_SIZE*, and 0 otherwise. All control variables presented in Table 4 are included in the models in this table, but are not reported for brevity.
See Appendix A for definitions of control variables.

with audit failures have larger abnormal accruals as evidenced by the positive and significant coefficients on *AUD_FAIL_10* ($p < 0.05$ and $p < 0.01$, two-tailed, respectively). In contrast, the F-tests indicate no evidence of contagion for larger offices. The results in Table 5 are consistent with prior studies that find that larger offices conduct audits of higher-quality (Francis and Yu 2009; Choi et al. 2010).[22] It also appears that the insignificant results for absolute abnormal accruals of non-Big 4 auditor clients in Table 4, Panel A are driven by the largest offices because the set of smaller non-Big 4 offices in Table 5 clearly show a contagion effect with respect to both absolute and income-increasing abnormal accruals.

[22] As an alternative to creating an indicator variable that splits offices into small and large, we use the continuous variable *OFFICE_SIZE* and interact this with *AUD_FAIL_X*. In a separate model we also interact *B4* with all other control variables. The interaction coefficients in all models as presented in Table 5 are negative and significant ($p < 0.05$), indicating that the contagion effect diminishes as office size increases.

Tests of H3—Contagion Effect and Big 4 Office-Level Industry Expertise

Table 6 presents the results of testing if a contagion effect in a Big 4 office depends on the percentage of audits performed in its areas of city-level industry leadership (H3). For brevity, Table 6 only reports the tests of income-increasing abnormal accruals at the 10 percent materiality threshold, but the results are similar when testing absolute abnormal accruals and when testing both types of accruals at the zero percent threshold. The presentation of included control variables is, again, suppressed.

The indicator variable *HIGH_OFFICE_EXPERTISE* in Table 6 is coded 1 when the value of *OFFICE_EXP_%* for a Big 4 office is greater than the sample median of 0.740.[23] The model specification also includes *AUD_FAIL_10*, and the interaction term *AUD_FAIL_10 ∗ HIGH_OFFICE_EXPERTISE*. Given this specification, the variable *AUD_FAIL_10* by itself tests if a contagion effect occurs in Big 4 offices in which the office conducts less than 74 percent of audits in its areas of city-level of industry leadership. The interaction term (*AUD_FAIL_10 ∗ HIGH_OFFICE_EXPERTISE*) tests if the results are incrementally different for auditor offices with high expertise, i.e., 74 percent or more of its audits are in its areas of city-level industry leadership. The F-statistic on the sum of the coefficients (*AUD_FAIL_10 + AUD_FAIL_10 ∗ HIGH_OFFICE_EXPERTISE*) tests the total contagion effect for offices where the auditor conducts a relatively *high* percentage of audits in its areas of city-level of industry leadership.

The first model in Table 6 analyzes all Big 4 offices. The positive and significant ($p < 0.01$) coefficient on *AUD_FAIL_10* indicates that a contagion effect exists in the Big 4 auditor offices that conduct a lower percentage of audits in its areas of industry expertise. In contrast, the F-statistic on the sum of *AUD_FAIL_10* and *AUD_FAIL_10 ∗ HIGH_OFFICE_EXPERTISE* is not significant, indicating no contagion for offices that conduct a larger percentage of audits in their areas of industry expertise. Thus, H3 is supported.

The second model in Table 6 examines just the largest quartile of Big 4 offices, and tests if contagion occurs in large offices that have a lower percentage of audits in the office's areas of industry leadership. Using the same model specification, none of the test variables are statistically significant at the 0.10 level, which means that the level of industry expertise has no effect for larger offices. Taken together, the results in Table 6 indicate that the use of industry expertise is important for the smaller 75 percent of Big 4 offices, but not for the largest quartile of office size. That is, the negative effect of smaller Big 4 office size, documented in Table 5, is mitigated when these offices conduct a larger percentage of audits in its areas of city-level industry leadership. In contrast, for the largest quartile of Big 4 offices, industry leadership has no statistical relation. An explanation is that the better human capital and/or quality control in the largest quartile of offices compensates for the relatively lower use of industry expertise. This finding extends Francis and Yu (2009) and Choi et al. (2010) by showing that a greater use of industry expertise mitigates the small-office effect documented in these studies.[24]

---

[23] We exclude all offices with less than ten clients in the analyses in Table 6 to avoid the possibility that very small offices are significantly influencing the calculation of *OFFICE_EXP_%*. This is possible because a very small denominator (total number of clients) for this variable calculation could create very large expertise percentages and possibly bias results.

[24] Similar to Table 5, as an alternative to creating an indicator variable that splits offices into low and high use of industry expertise, we use the continuous variable *OFFICE_EXP_%* and interact this with *AUD_FAIL_X*. In a separate model we also interact *B4* with all other control variables. The interaction coefficient in model one is negative and significant ($p < 0.10$) indicating that the contagion effect diminishes as offices make use of more industry expertise. The interaction coefficient is insignificant in model two, which is consistent with our tabled results.

TABLE 6

Audit Failures in Big 4 Offices and Income-Increasing Abnormal Accruals Dependent on City-Level Industry Expertise

		Dependent Variable is *ABN_ACC* > 0	
Variable	**Pred.**	**All Big 4 Offices**	**Largest Big 4 Offices**
Test Variables			
HIGH_OFFICE_EXPERTISE	?	−0.005	−0.013
AUD_FAIL_10	+	0.010***	0.001
(p-value)		(0.003)	(0.875)
*AUD_FAIL_10 * HIGH_OFFICE_EXPERTISE*	−	−0.005	0.001
(p-value)		(0.209)	(0.856)
F-stat [*AUD_FAIL_10 + AUD_FAIL_10 * HIGH_OFFICE_EXPERTISE*]		2.1	0.1
(p-value)		(0.172)	(0.720)
All Control Variables		Yes	Yes
Year Fixed Effects		Yes	Yes
Industry Fixed Effects		Yes	Yes
SEC Regional Office Fixed Effects		Yes	Yes
n		5,664	3,821
Model p-value		< 0.001	< 0.001
R^2		33.0%	30.8%

*, **, *** Indicate significance at the 0.10, 0.05 and 0.01, respectively, using two-tailed tests.
t-statistics are calculated based on robust standards errors clustered at the auditor-office level. The "Largest Big 4 Offices" are Big 4 auditor offices that are larger than the 75th percentile value of *OFFICE_SIZE*, and 0 otherwise. The 75th percentile value for Big 4 offices is $4.9 million in audit fees. *ABN_ACC* is the signed value of a company's abnormal accruals as calculated per Kothari et al. (2005). Only observations with values of *ABN_ACC* greater than 0 (income-increasing abnormal accruals) are analyzed. *AUD_FAIL_10* is 1 when at least one audit failure occurs within the same office of a company's external auditor during year *t*, and 0 otherwise. An audit failure is defined as existing when at least one of the auditor's client companies within an auditor office subsequently restates net income downward in the future by 10 percent or more of originally reported net income. *HIGH_OFFICE_EXPERTISE* is 1 when the percentage of audits conducted within an auditor office in a year where the auditor is the city-level industry expert is greater than the median value of 0.740 for all Big 4 auditor-office-year observations, and 0 when it is below the median value. The percentage is calculated for each office-year by scaling the number of audits conducted within an office in a year where the auditor is the city-level industry expert by the total number of audit engagements in the office in the same year. An auditor is the city-level industry expert when it has the highest market share of audit fees within a Metropolitan Statistical Area in an industry defined by two-digit SIC codes. All control variables presented in Table 4 are included in the models in this table, but are not reported for brevity.
See Appendix A for definitions of control variables.

Sensitivity Analyses

The results are robust to an alternative coding of the office-year test variable *AUD_FAIL_X*, which takes into consideration the size of the office. Specifically, we use the variable *PERC_FAIL_X*, which is calculated as the number of downward restatements in an office-year scaled by the number of clients in the office-year. Therefore, this variable reflects the percentage of clients that restate in an auditor office-year. The results in Tables 4 through 6 are robust to this alternative coding of the office-year test variable (note this analysis is not pertinent to Table 3). More specifically, results are virtually identical to those tabulated if we use a ranked version of *PERC_FAIL_X*, where *PERC_FAIL_X* is sorted into 11 ranks (the ranked variable equals 0 if there

are zero audit failures in an office-year, and is then ranked into 10 additional partitions where the number of audit failures in an office-year is greater than 0, with an equal number of observations within each partition). Results are qualitatively similar if we use the raw form of *PERC_FAIL_X* instead of the ranked form.

We also investigate whether our results are robust to controlling for whether a company is an accelerated filer, and controlling for the quality of client companies' internal controls. We do this by creating two new variables: *ACCEL*, which equals 1 for firm-year observations where the company is an accelerated filer, and *SOX404FAIL*, which equals 1 when the audit report deems the company's internal controls as ineffective. Given that auditors began issuing audit opinions regarding internal controls only after Sarbanes-Oxley, and that data on this variable appear in Audit Analytics beginning only in 2004, we code *SOX404FAIL* as equal to 1 for a company in the years 2000 through 2003 if *SOX404FAIL* is equal to 1 in 2004. We do this to retain our full sample because it seems reasonable to assume that if a company's internal controls were ineffective in 2004 it is very likely they were also ineffective in the recent years prior to 2004. In all analyses *ACCEL* is always negative, but is significant in only about a quarter of the regression tests. Further, *SOX404FAIL* is never significant. Finally, and most importantly, our results in Tables 3 through 6 are not affected by the addition of these two variables and our results remain very similar to those tabulated.[25]

We also test whether our results are robust to calculating materiality with respect to originally reported sales rather than to net income. Specifically, we scale the downward change in net income by a company's net sales and consider materiality levels of 1, 2, and 10 percent of net sales to be an audit failure. All tabulated results in the study are very similar using all three sales-based materiality thresholds.

As an additional control, the models are re-estimated with auditor-office fixed effects, and these results are qualitatively similar to those reported in the tables. This procedure controls for potential omitted correlated variables with respect to office characteristics, and gives additional confidence the results in the study are not the consequence of omitted office variables.

We also investigate if one or more specific Big 4 audit firms is systematically driving the results, or whether the contagion effect is a characteristic of all of the Big 4 firms. To do so, we rerun all of our regression models in Tables 3 through 6 on separate samples of each individual Big 4 auditor. These auditor-specific regressions are qualitatively the same as the pooled results roughly two-thirds of the time. Further, differing results are spread evenly among the individual Big 4 audit firms. We conclude that there exists no systematic difference in the contagion effect across the Big 4 firms.

Many non-Big 4 accounting firms have only a few office locations. Untabulated analysis shows that out of the total of 1,997 non-Big 4 offices in the sample, 794 are single-office audit firms. Therefore, it is possible that these very small audit firms are driving the results for non-Big 4 auditors. In order to test whether the main results are sensitive to this we delete all observations where a company is audited by a non-Big 4 audit firm that has only one office location. All tabulated results are very similar using this reduced sample.

Similarly, it is possible that auditor offices with very few clients are driving the results. To address this concern we delete all companies audited by a non-Big 4 (Big 4) auditor office where the total number of clients of the office in a year is less than two (four). These client numbers represent the 25th percentile level for each auditor type (i.e., 25 percent of non-Big 4 and Big 4

---

[25] The *SOX404FAIL* variable used in the re-analyses of models in Table 3 is the mean value of this variable for all clients within an office-year.

auditor offices have less than two and four clients, respectively). Results are very similar compared to tabulated results.

Audit Analytics (AA) identifies a subset of firm-level restatements as occurring specifically due to accounting irregularities. As a sensitivity test we consider only downward restatements of net income due to these irregularities to be an audit failure. All results are qualitatively the same for these restatements compared to the tabulated results in terms of coefficient signs and statistical significance.

Finally, a concurrent study by Francis et al. (2012) makes a limited comparison of income restatements in the Audit Analytics and Compustat Unrestated Quarterly databases and finds that both contain some errors, but when the two databases are in agreement, they correspond to SEC filings. Consequently, we perform a robustness test by keeping only restatement observations where Audit Analytics and Compustat are in agreement. Since Audit Analytics reports only the cumulative amount of restatements, we average this over the restatement period to estimate the yearly effects that are then compared with Compustat. We retain only those restatements where the office-year test variable *AUD_FAIL_X* is coded the same using both databases. For this reduced sample, untabulated results indicate that the tests are virtually the same as those tabulated for Big 4 offices. Specifically, all Big 4 results in Tables 3 through 6 are the same in terms of the signs on the coefficients and significance levels for *AUD_FAIL_X* except for one model where significance is at the 0.10 level compared to the tabulated 0.05 level. Results for non-Big 4 offices are qualitatively the same for about three-quarters of the models in Tables 3 through 6. We conclude that potential measurement error in the restatement databases does not affect the study's results.

V. CONCLUSION

We investigate if the presence of an audit failure in an office-year reveals the likelihood of a "contagion effect" that results in other concurrent low-quality audits in the office. We test this by determining if the presence of at least one audit failure, measured as the downward restatement of client earnings, indicates a systematic problem in the average quality of concurrent audits performed in the same office for the same year that the misstated earnings were originally reported. We first show that offices with client restatements in the past are more likely to have new client restatements for up to five years in the future, suggesting a contagion of audit failures over time, and that accounting firms do not seem to quickly identify and remediate the quality-control problems in these offices.

Next we examine if the earnings quality (abnormal accruals) of clients in offices with audit failures is lower, on average, than clients in those offices with no audit failures. We find that in auditor offices where an audit failure occurred, the concurrent clients of that office (in the same year) have a higher level of abnormal accruals compared to offices with zero audit failures. These results hold for all but the largest quartile of office size, which indicates that office size is also an important factor in contagion. Last, we find that a relatively high use of industry expertise in a Big 4 office can mitigate the negative effect of small office size.

As discussed in the introduction, our findings should be of interest to regulators, audit standard-setters, accounting firms, and investors because we provide a method to infer the overall quality of an auditor office location through the use of easily obtainable and publicly available information on restatements. Regulators can use this information to identify offices where audits are more likely to be of lower quality (and perhaps even below minimum standards), and standard-setters may be able to use this information to develop standards that emphasize the potential for quality-control problems in the offices of multi-location audit firms. Accounting firms can benefit from this method as those in charge of quality-control processes in a firm's national office can use it to identify specific offices that may not be implementing the firm's quality-control

procedures appropriately. Finally, investors may be able to use it as one piece of additional information with which to infer something about the earnings quality of a particular company based on the history of the auditor office that performs the audit.

The study also has some limitations. First, our measure of an audit failure within an auditor office relies on the assumption that every company that "should" restate earnings actually does so. This is not likely to be the case. However, we note that in cases where an audit engagement is misclassified as not being an audit failure (no restatement is issued) when it may in fact be a poor-quality audit (that is unobservable) biases against finding a statistically significant result. Second, our sample is limited to audits of publicly traded companies. To the extent that a particular auditor office also performs audits of smaller, private companies, our results cannot be generalized to this client base.

REFERENCES

Altman, E. 1983. *Corporate Financial Distress: A Complete Guide to Predicting, Avoiding, and Dealing With Bankruptcy*. New York, NY: Wiley.

American Institute of Certified Public Accountants (AICPA). 1972. *Responsibilities and Functions of the Independent Auditor*. Statement on Auditing Standards No. 1, Section 110. New York, NY: AICPA.

Antle, R., and B. Nalebuff. 1991. Conservatism and auditor-client negotiations. *Journal of Accounting Research* 29 (Supplement): 31–54.

Basu, S. 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics* 24: 3–37.

Becker, C., M. DeFond, J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* (Spring): 1–24.

Bronson, S. N., C. E. Hogan, M. F. Johnson, and K. Ramesh. 2011. The unintended consequences of PCAOB Auditing Standards Nos. 2 and 3 on the reliability of preliminary earnings releases. *Journal of Accounting and Economics* 51 (1-2): 95–114.

Cheffers, M., D. Whalen, and O. Usvyatsky. 2010. 2009 financial restatements: A nine year comparison. *Audit Analytics Sales* (February).

Choi, J. H., C. F. Kim, J. B. Kim, and Y. Zang. 2010. Audit office size, audit quality, and audit pricing. *Auditing: A Journal of Practice & Theory* 29 (1): 73–97.

Comprix, J., L. Mills, and A. Schmidt. 2012. Bias in quarterly estimates of annual effective tax rates and earnings. *Journal of the American Taxation Association* 34 (1): 31–53.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

Francis, J. R. 2011. A framework for understanding and researching audit quality. *Auditing: A Journal of Practice & Theory* 30 (2): 125–152.

Francis, J. R., and J. Krishnan. 1999. Accounting accruals and auditor reporting conservatism. *Contemporary Accounting Research* 16 (1): 135–165.

Francis, J. R., and M. Yu. 2009. Big 4 office size and audit quality. *The Accounting Review* 84 (5): 1521–1553.

Francis, J. R., E. L. Maydew, and H. C. Sparks. 1999a. The role of Big 6 auditors in the credible reporting of accruals. *Auditing: A Journal of Practice & Theory* 18 (2): 17–34.

Francis, J. R., D. Stokes, and D. Anderson. 1999b. City markets as a unit of analysis in audit research and the re-examination of Big 6 market shares. *Abacus* 35: 185–206.

Francis, J. R., K. Reichelt, and D. Wang. 2005. The pricing of national- and city-specific reputations for industry expertise in the U.S. audit market. *The Accounting Review* 80 (1): 113–136.

Francis, J. R., P. Michas, and M. Yu. 2012. *Office Size of Big 4 Auditors and Client Restatements. Contemporary Accounting Research* (forthcoming).

Frankel, R., M. Johnson, and K. Nelson. 2002. The relation between auditors' fees for nonaudit services and earnings management. *The Accounting Review* 77 (Supplement): 71–105.

Gleason, C., N. Jenkins, and W. B. Johnson. 2008. The contagion effects of accounting restatements. *The Accounting Review* 83 (1): 83–110.
Hiltebeitel, K. M., B. A. Leauby, and J. M. Larkin. 2000. Job satisfaction among entry-level accountants. *The CPA Journal* LXX (5): 76–78.
Hribar, P., and C. Nichols. 2007. The use of unsigned earnings quality measures in tests of earnings management. *Journal of Accounting Research* 45 (5): 1017–1053.
Johnson, V., I. K. Khurana, and J. K. Reynolds. 2002. Audit firm tenure and the quality of accounting earnings. *Contemporary Accounting Research* 19:637–660.
Johnstone, K. M., and J. C. Bedard. 2004. Audit firm portfolio management decisions. *Journal of Accounting Research* 42 (4): 659–690.
Jones, J. J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (2): 193–228.
Kedia, S., and S. Rajgopal. 2011. Do the SEC's enforcement preferences affect corporate misconduct? *Journal of Accounting and Economics* 51 (3): 259–278.
Kennedy, P. 1992. *A Guide to Econometrics*. Cambridge, MA: MIT Press.
Kinney, W. R., Z.-V. Palmrose, and S. Scholz. 2004. Auditor independence, non-audit services, and restatements: Was the U.S. government right? *Journal of Accounting Research* 42 (3): 561–588.
Kothari, S. P., A. J. Leone, and C. E. Wasley. 2005. Performance matched discretionary accruals. *Journal of Accounting and Economics* 39 (1): 163–197.
Kothari, S. P., T. Lys, C. W. Smith, and R. L. Watts. 1989. Auditor liability and information disclosure. *Journal of Accounting, Auditing and Finance* 4: 307–339.
Krishnan, G. V. 2005. Did Houston clients of Arthur Andersen recognize publicly available news in a timely fashion? *Contemporary Accounting Research* 22 (1): 165–193.
Palmrose, Z.-V., and S. Scholz. 2004. The circumstances and legal consequences of non-GAAP reporting: Evidence from restatements. *Contemporary Accounting Research* 21 (1): 139–180.
Panel on Audit Effectiveness. 2000. Report and recommendations: Public oversight board of the American Institute of Certified Public Accountants. Available at: http://www.pobauditpanel.org/download.html
Plumlee, M., and T. L. Yohn. 2010. An analysis of the underlying causes attributed to restatements. *Accounting Horizons* 24 (1): 41–64.
Price R. A., III, N. Y. Sharp, and D. A. Wood. 2011. Detecting and predicting accounting irregularities: A comparison of commercial and academic risk measures. *Accounting Horizons* 24 (4): 755–780.
Reichelt, K. J., and D. Wang. 2010. National and office-specific measures of auditor industry expertise and effects on audit quality. *Journal of Accounting Research* 48: 647–686.
Reynolds, J. K., and J. R. Francis. 2000. Does size matter? The influence of large clients on office-level auditor reporting decisions. *Journal of Accounting and Economics* 30:375–400.
Skinner, D. J. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32: 38–60.
Wallman, S. 1996. The future of accounting, Part III: Reliability and auditor independence. *Accounting Horizons* 10: 76–97.

APPENDIX A
VARIABLE DEFINITIONS

Dependent Variables

ABS_ABN_ACC = absolute value of a company's abnormal accruals as calculated in Kothari et al. (2005), with control for concurrent return on assets; and

ABN_ACC = signed value of a company's abnormal accruals as calculated in Kothari et al. (2005). Analyses using *ABN_ACC* examine only income-increasing abnormal accruals.

Test Variables

B4 = 1 if the company hires a Big 4 auditor in year *t*, and 0 otherwise;

AUD_FAIL_X = 1 when at least one low-quality audit failure occurs within the same office of a company's external auditor during year *t*, and 0 otherwise. A low-quality audit is defined as existing when a client company restates net income downward by a material amount subsequent to the audit. *X* refers to the materiality level of the restatement (i.e., 0 for a greater than 0 percent downward restatement of net income and 10 for a greater than 10 percent downward restatement of net income). Auditor office locations are taken from Audit Analytics;

LARGE_OFFICE = 1 for Big 4 auditor offices that are larger than the 75th percentile value of *OFFICE_SIZE* (see below for the definition of *OFFICE_SIZE*), and 0 otherwise. The 75th percentile value for non-Big 4 (Big 4) offices is $667,000 ($13.5 million) in audit fees. This variable is used only in Tables 5 and 6; and

HIGH_OFFICE_EXPERTISE = 1 when the percentage of audits conducted within an auditor office in a year where the auditor is the city-level industry expert is greater than the median value of 0.740 for all Big 4 auditor-office-year observations, and 0 when it is below the median value. The percentage is based on the definition of *OFFICE_EXP_%*, which is defined below. An auditor is the city-level industry expert when it has the highest market share of audit fees within a Metropolitan Statistical Area in an industry defined by two-digit SIC codes (see the variable definition for *CITY_IND_EXP* below). This variable is used only in Table 6.

Auditor Office-Level Control Variables

OFFICE_SIZE = natural log of the total dollar amount of audit fees charged to all audit clients within an auditor office in year *t*. Auditor office locations are taken from Audit Analytics;

RISK_PORT = mean value of the average of client assets, leverage, and return on assets for an auditor office in year *t*. The average is calculated by taking the mean level of client assets, leverage, and return on assets. These means are then standardized so that each has a mean of 0 and a standard deviation of 1. These standardized values are then averaged together to form the overall client risk portfolio for the auditor office in a year; and

OFFICE_EXP_% = number of audits conducted in a Bog 4 office in a year where that office is the city-level industry leader, scaled by the total number of audits conducted by the office in the year.

Firm-Level Control Variables

CITY_IND_EXP = 1 if the company's auditor is the city-level industry expert auditor, and 0 otherwise, where industry expertise is calculated based on total audit fees charged by the audit firm to clients within a particular Metropolitan Statistical Area and industry (similar to Francis et al. 2005). The audit firm with the highest amount of audit fees within an industry in a city-year is classified as the city-level industry expert. Industries are defined at the two-digit SIC code level;

NAT_IND_EXP = 1 if the company's auditor is the national-level industry expert auditor, and 0 otherwise, where industry expertise is calculated based on total audit fees charged by the audit firm to clients in a particular industry within the U.S. The audit firm with the highest amount of audit fees within an industry-year is classified as the national industry expert. Industries are defined at the two-digit SIC code level;

INFLUENCE = total dollar value of both audit and nonaudit fees charged to a specific client in year *t*, scaled by the total audit fees charged by the auditor office in the same year;

SIZE = natural log of a company's total assets in year *t*;

LAG_TOT_ACC = company's total accruals scaled by total assets in year t–1;
CFO = company's cash flows from operations in year t scaled by lagged total assets;
CFO_VOL = standard deviation of a company's cash flows from operations from year t–2 through year t;
SALES_GROWTH = one-year percentage growth in a company's sales from year t–1 to year t;
SALES_VOL = standard deviation of a company's sales from year t–2 through year t;
PPE_GROWTH = one-year percentage growth in a company's net property, plant, and equipment from year t–1 to year t;
LEV = company's total debt in year t, scaled by lagged total assets;
MB = company's market value of equity scaled by book value of equity at the end of year t;
RETURN = company's 12-month stock return during year t–1;
RET_VOL = standard deviation of a company's monthly stock returns during year t;
SHARE_ISSUE = 1 if the company issues additional shares in year t, and 0 otherwise;
LOSS = 1 if the company records net income below 0 in year t, and 0 otherwise;
LITIGATE = 1 if a company is within the following SIC codes: 2833–2836, 3570–3577, 3600–3674, 5200–5961, and 7370), and 0 otherwise;
BANKRUPTCY = probability of bankruptcy using the Altman-Z score [(0.717 * net working capital/assets) + (0.847 * retained earnings/assets) + (3.107 * earnings before interest and taxes/assets) + (0.42 * book value of equity/liabilities) + (0.998 * sales/assets)] (Altman 1983);
#_OPER_SEGS = number of operating segments the company operates in during year t; and
#_GEO_SEGS = number of geographic segments the company operates in during year t.

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 553–575

American Accounting Association
DOI: 10.2308/accr-50312

The Effect of Relative Performance Information on Performance and Effort Allocation in a Multi-Task Environment

R. Lynn Hannan
Georgia State University and Center for the Economic Analysis of Risk

Gregory P. McPhee
Georgia State University

Andrew H. Newman
University of Pittsburgh

Ivo D. Tafkov
Georgia State University

ABSTRACT: This study investigates how relative performance information (RPI) affects employee performance and allocation of effort across tasks in a multi-task environment. Based on behavioral theories, we predict that the social comparison process inherent in RPI induces both a motivation effect that results in increased effort as well as an effort distortion effect that results in the distortion of effort allocations across tasks away from the firm-preferred allocations. We also predict that both effects are magnified when the RPI is public compared to private. We argue that although the motivation effect will generally benefit performance, the effort distortion effect may be detrimental to performance. We design an experiment that isolates these two effects. Consistent with our predictions, we find that RPI induces both motivation and effort distortion effects and that both effects are magnified when the RPI is public rather than private. Although the motivation effect increases performance, we demonstrate that the effort distortion effect can decrease performance. By isolating the motivation and effort distortion effects, our study provides insights into the costs and benefits of RPI in a multi-task environment. As

We thank Willie Choi, Bill Dilla, Harry Evans (senior editor), Gary Hecht, Steve Kachelmeier, Kathryn Kadous, Ranjani Krishnan, Joan Luft, Victor Maas, Laureen Maines (editor), Lil Mills, Tim Mitchell, Anthony Nikias, Karen Sedatole, Mike Shields, Bill Tayler, Kristy Towry, Marcel van Rinsum, Michael Williamson, two anonymous reviewers; and participants at Erasmus University, Michigan State University, and The University of Texas at Austin workshops, Emory University and Georgia State University brownbags, the 2010 *Accounting, Behavior and Organizations* Research Conference, the 2011 Management Accounting Section Research and Case Conference, and the 2011 American Accounting Association Annual Meeting for their helpful comments. We also appreciate the financial support provided by the Institute of Management Accountants Foundation for Applied Research and Georgia State University.

Editor's note: Accepted by Laureen Maines.

Submitted: March 2011
Accepted: September 2012
Published Online: September 2012

such, it informs accountants regarding the design of information systems and when tasks should be aggregated or disaggregated across employees.

Keywords: *multiple tasks; relative performance; social comparison theory.*

Data Availability: *Data are available from the authors upon request.*

I. INTRODUCTION

This paper reports the results of an experiment that investigates the behavioral effects of relative performance information (RPI) on performance and the allocation of effort in a multi-task environment. Specifically, we investigate how private RPI and public RPI affect employee performance and the allocation of effort across tasks. We isolate the behavioral effects of RPI by using a setting in which employees are not compensated based on performance output.[1]

As explained below, our setting is intended to capture a multi-task environment where the firm has preferences over the allocation of effort across tasks, communicates these preferences to employees, and pays the employees flat wages. We modify this setting in our experiment by compensating participants based on effort allocations rather than a flat-wage contract. This modification allows us to clearly attribute deviations from firm-preferred allocations to behavioral factors. Such deviations are referred to as distorted effort allocations in this paper and represent deviations from firm preferences and not necessarily from an efficient contract.

A multi-task environment is one in which an employee allocates effort across either multiple dimensions of a single task or multiple tasks within a single job description. Many job tasks are inherently multi-dimensional. For example, employees working on production tasks are responsible for both the quantity and the quality of their output, and employees working in a hotel reservation center are evaluated on both the number of calls they convert to bookings and the revenue generated per booking (Berger et al. 2011). Firms may also combine multiple tasks into a single job in order to exploit synergies (Lindbeck and Snower 2000). For example, customer service employees in the financial services industry have the unique opportunity to inform customers about the additional services offered by the firm while responding to customers' requests. As such, these employees are responsible for both providing traditional customer service and cross-selling other products to customers (Askin and Harker 1999). Similarly, many hospital-based physicians are responsible for examining patients, performing administrative duties, teaching medical students, and conducting research (Dumont et al. 2007).

Importantly, in a multi-task environment, effort expended toward one task cannot be simultaneously expended toward another. Therefore, how an employee allocates effort can have important consequences for the performance of the employee and the firm. Analytic models have investigated the complexities inherent in designing an appropriate incentive system in the multi-task environment. Such complexities arise because the design must consider not only how to motivate the desired level of effort, but also the appropriate allocation of effort across tasks (Holmstrom and Milgrom 1991; Feltham and Xie 1994; Datar et al. 2001). Consistent with this perspective, our study investigates how RPI affects both effort level (i.e., motivation) and effort allocation. Understanding these effects can provide important insights to the firm regarding the costs and benefits associated not only with providing RPI, but also with aggregating tasks into a single job.

---

[1] We do not provide performance-based incentives because the primary focus of our study is identifying the behavioral effects of RPI in multi-task environments. Non-performance based compensation is commonly observed in the workplace (Baker et al. 1988), which enhances the generalizability of our results.

Employees may obtain RPI via informal methods, such as by observing peers who work in close proximity or formal information systems provided by the firm. Firms commonly provide RPI to their employees even when it is not linked to compensation (e.g., Wikoff et al. 1982; Anderson et al. 1982; Nordstrom et al. 1990). In banking, for example, some branch managers disclose to their tellers the number of new accounts opened by each of their colleagues and the total amount of funds in these accounts even though the tellers' compensation is not based on peer performance (Tafkov 2011). Retail stores display individual and relative performance metrics such as average sales per hour, units sold, and dollars per transaction when salespersons log onto their terminals and base scheduling, but not pay, on cashiers' relative performance on these metrics (O'Connell 2008). More recently, firms have turned to "gamification" as a technique for motivating the workforce by rewarding employees with points for completing assignments and using leader boards so that employees can view one another's scores (Silverman 2011).

One of the important functions of accountants is determining how information affects employee behavior. Prior research has investigated how RPI affects employee behavior in single-task environments and found that it can improve performance with flat wage and individual piece rate incentive schemes (Kerr et al. 2007; Hannan et al. 2008; Tafkov 2011), although not necessarily with tournament incentive schemes (Hannan et al. 2008; Newman and Tafkov 2011).[2] The theory behind the improved performance findings is that RPI encourages social comparison, which in turn motivates higher effort and performance. We call this the motivation effect of RPI. We predict that RPI in the multi-task environment induces motivation effects similar to the single-task environment and, therefore, is potentially beneficial to the firm. Consistent with Tafkov's (2011) findings in a single-task environment, we also predict that the motivation effect of RPI is magnified when the RPI is public compared to private.

In the multi-task environment, RPI may also induce an effort distortion effect. That is, employees may distort their effort allocations away from firm-preferred proportions in order to do well on some tasks even if it means that they do less well on other tasks. The theory behind this prediction is that doing so will ward off a threat to their self-image in one area by affirming their competence in another area (Steele 1988) and/or satisfy an innate desire for social distinction (Frey 2007) by outperforming peers in at least one area. We also predict that the effort distortion effect of RPI is magnified when the RPI is public compared to private. Whereas the motivation effect benefits the firm, the effort distortion effect may decrease performance and, therefore, harm the firm. Given these potentially offsetting effects, we investigate the effect of RPI on performance in the multi-task environment via research questions.

We investigate the behavioral effects of RPI on performance and effort allocation in a multi-task environment via an experiment that uses a 2×3 between-subjects design. We vary whether participants can choose how much effort to allocate between two tasks (no choice, choice) in order to isolate the motivation effect from the effort distortion effect of RPI. We also vary RPI at three levels: *No RPI*, *Private RPI* (i.e., each individual's relative performance rank is known only by that individual), and *Public RPI* (i.e., each individual's relative performance rank is known by that individual and all peers). Although participants in the *Private RPI* and *Public RPI* conditions receive RPI, no participant's compensation is based on performance. Rather, participants are compensated based on whether their effort allocations between tasks are consistent with firm-preferred allocations. For simplicity, we set the allocation as equal across the two tasks. The two tasks are structured to reflect the diminishing marginal returns to effort that is inherent in many tasks. Our primary dependent variables are *Time Difference*, which measures the degree of

[2] Frederickson (1992) also investigates how RPI affects effort in a single-task environment in which he holds RPI constant while manipulating common uncertainty and incentive scheme.

effort distortion away from participants' wealth-maximizing equal allocation of effort, and *Total Problems Solved*, which measures the overall performance output on the two problem-solving tasks.

Consistent with our predictions, we find that private RPI and public RPI induce a motivation effect and an effort distortion effect. Both effects are magnified for public RPI. Specifically, when participants have no choice in effort allocations, and thus only the motivation effect can influence performance, private RPI increases performance and public RPI increases it to a greater extent. When participants can choose their effort allocations, they distort their allocations away from an equal allocation of effort only when they receive private RPI or public RPI; they do not distort their allocations in the absence of RPI. Further, effort allocations are more distorted for public RPI compared to private RPI. Finally, we find that the effort distortion effect is large enough to offset the positive motivation effect of RPI on performance only when the RPI is public. That is, we demonstrate that in our experimental setting, the effort distortion effect caused by public RPI is detrimental to performance.

The results of our study have important implications for both practice and theory. Our study informs accountants regarding the design of information systems as well as the assignment of task responsibilities. Specifically, when the firm has the ability to provide RPI to its employees, our results suggest that the decision to provide RPI may depend on whether the firm is able to control how employees allocate their effort across tasks. If the firm is able to control effort allocations, then providing RPI, and especially public RPI, is likely to have a positive effect on performance because only the motivation effect is present.

When employees have discretion over how they allocate their effort across tasks, however, the RPI-induced effort distortion effect makes it more difficult to predict the impact of RPI on performance. Our study demonstrates that the effort distortion effect can negatively affect performance in environments where tasks have diminishing marginal returns to effort. Although this is likely to be descriptive of the tasks in many workplace environments, some environments may involve tasks with increasing marginal returns to effort, suggesting that the effort distortion effect could increase performance. Our study suggests, therefore, that when employees have discretion over how they allocate effort across tasks, accountants should consider the marginal returns to effort when determining whether to provide RPI. Our study also suggests that if RPI is readily available via informal sources, then accountants should consider the potential effects of distorted effort allocations when determining whether to aggregate tasks into single jobs.

By identifying the effects of RPI in the multi-task environment, our study contributes to two streams of research. First, our study contributes to the stream of research investigating effort allocation concerns in multi-task and multi-attribute environments. Although prior studies have not considered the effect of RPI, both analytic and experimental studies (Holmstrom and Milgrom 1991; Feltham and Xie 1994; Datar et al. 2001; Farrell et al. 2008; Kachelmeier et al. 2008; Kachelmeier and Williamson 2010) have investigated the design of incentive contracts in multi-task environments. Other studies have investigated how social preferences such as fairness (Fehr and Schmidt 2004) or social pressure (Bruggen and Moers 2007) may influence the allocation of effort. Our study adds to this growing multi-task literature. Second, prior research has documented the positive motivation effect of RPI, and public RPI in particular, in single-task environments (Kerr et al. 2007; Hannan et al. 2008; Tafkov 2011). Our study provides evidence that this positive effect may be offset by an effort distortion effect in multi-task environments.

Section II next provides background information and develops the hypotheses; Section III describes the experimental design; Section IV reports the results; and Section V summarizes and discusses the results.

II. BACKGROUND AND HYPOTHESES

Inherent in the multi-task environment is the complexity of motivating not only the appropriate level of effort, but also the appropriate allocation of effort across tasks (Holmstrom and Milgrom 1991). We investigate the effects of RPI on effort level (i.e., motivation) and effort allocations in a setting where the employee's pay is not based on performance output, either individual or relative. According to conventional economic theory, RPI should have no effect on motivation or effort allocations in this setting because compensation is not based on relative performance (Kandel and Lazear 1992). As such, we rely on behavioral theories to develop our hypotheses. Social comparison theory (Festinger 1954) posits that individuals have a drive to compare themselves to others in order to evaluate their own abilities, which in turn affects their self-image. Individuals are motivated to maintain a positive self-image (Beach and Tesser 1995; Tesser 1988). Performing better than others leads to positive feelings such as pride, while performing worse than others leads to negative feelings such as shame (Lazarus 1991; Smith 2000). To generate positive feelings and maintain a positive self-image, individuals want their performance to be better than that of others, leading them to behave competitively when they are able to compare their performance to others (Festinger 1954; Hoffman et al. 1954).

We argue that the competitive behavior resulting from RPI manifests itself differently depending on the task environment in the workplace. In both single-task and multi-task environments, employees can attempt to perform better than their peers by exerting more effort. As a result, RPI may induce a motivation effect in both environments. In the multi-task environment, however, RPI may also induce an effort distortion effect whereby employees adjust their effort allocations away from firm-preferred proportions so as to do well on one task even if it means that they do less well on another task. Self-affirmation theory (Steele 1988) posits that individuals may cope with a threat to their self-image in one area by affirming their competence in another area. It follows, therefore, that employees may apply this coping mechanism in the workplace by shifting their effort toward tasks where they will be able to affirm their competence. Further, individuals have an innate desire to achieve positive social distinction (Frey 2007). To the extent that the effort shift allows employees to outperform peers in at least one area, it enhances their social distinction.

In the following sections we develop hypotheses to isolate the motivation effect of RPI from the effort distortion effect of RPI. We also predict that both of these effects are magnified when RPI is public compared to private.

Motivation Effect of RPI

We first consider the effect of RPI in a setting where employees have no discretion in effort allocations in order to isolate the motivation effect. Consistent with the tenets of social comparison theory, as described above, prior research finds that RPI-induced competitive behavior has a positive effect on performance in single-task environments when financial incentives are not tied to relative performance (Kerr et al. 2007; Hannan et al. 2008; Tafkov 2011). We expect to replicate these findings in the multi-task environment.

Consistent with Tafkov's (2011) findings in a single-task environment, we also predict that the motivation effect will be magnified when RPI is public compared to private. Private RPI involves employees knowing only their own relative standing compared to peers, and public RPI involves employees knowing not only their own relative standing, but also the relative standings of each of their peers. As discussed above, the social comparison process leads to feelings of pride if one's performance is above that of one's peers and shame if one's performance is below that of one's peers (Lazarus 1991; Smith 2000). These feelings are likely to be magnified if one's relative performance is public information. Smith (2000, 188) observes, "[J]ust as our inferior, blameworthy attributes create less shame if they are kept private, our superior, praiseworthy

attributes create greater pride if they are made public." Empirical research supports the notion that public comparison has a strong effect on comparison choices (Wilson and Benner 1971; Smith and Insko 1987; Tafkov 2011). Further, impression management theory posits that individuals are concerned with the impressions of others and will take actions to avoid negative impressions and create positive impressions (Goffman 1959; Tadeschi 1981). Public RPI introduces the added motivation to maintain a positive impression among one's peers. Therefore, we expect to replicate Tafkov's (2011) findings that, in the single-task environment, RPI has a positive effect on performance and that this effect is greater when RPI is public compared to private. Our hypothesis is:

H1: When employees cannot alter their effort allocations across tasks, employees with public relative performance information will perform better than employees with private relative performance information, and employees with private relative performance information will perform better than employees with no relative performance information.

Effort Distortion Effect of RPI

We now consider the effect of RPI in a setting where employees have discretion over how to allocate their effort across tasks. Although employees may distort their effort allocations away from firm-preferred proportions even in the absence of RPI (e.g., toward tasks employees find intrinsically appealing), RPI can induce a shift toward greater distortion. As discussed above, such a shift may occur because individuals can address a threat to their self-image in one area by affirming their competence in another area (Steele 1988). It may also occur because the innate desire to distinguish themselves positively relative to others (Frey 2007) may motivate individuals to seek such distinction by performing better than their peers on at least one task. Individuals strive to be different from others in a positive way (Snyder and Fromkin 1980; Lynn and Snyder 2002) and such positive differences enhance self-image and social status (Tesser 1988; Lynn and Snyder 2002; Simsek and Yalincetin 2010).

The desire for self-affirmation and social distinction may cause employees to be willing to reduce their performance on some tasks in order to boost their performance on other tasks. Consistent with this perspective, Brickman and Bulman (1977) find that individuals are more satisfied when they receive feedback that they performed very well on one task and very poorly on another task compared to when they receive feedback that their performance was mediocre on both tasks. To the extent that the desire for self-affirmation and social distinction motivates employees to perform better on at least one task, RPI will induce distorted effort allocations.

The above discussion leads to our prediction that when employees have discretion over how to allocate their effort across tasks, their effort allocations will be more distorted, i.e., less compliant with firm-preferred proportions, when they have RPI compared to when they do not. Consistent with the theory presented in relation to H1, we also expect that public RPI will induce a greater motivation for employees to perform well on at least one task in order to exhibit a positive and/or socially distinguished image to their peers. As a result, employees will distort their effort allocations to a greater extent when they have public RPI compared to private RPI. Our hypothesis is:

H2: When employees can alter their effort allocations across tasks, employees with public relative performance information will distort their effort allocations across tasks more than employees with private relative performance information, and employees with private relative performance information will distort their effort allocations more than employees with no relative performance information.

We now consider how RPI affects performance output when employees can alter their effort allocations. Although H2 predicts that RPI will lead to distorted effort allocations, we cannot predict how such distortions impact performance because it depends on the relative marginal returns to effort of tasks that are emphasized compared to those that are de-emphasized. That is, if an employee distorts effort toward tasks with greater marginal returns to effort, performance will increase. On the other hand, if an employee distorts effort toward tasks with lower marginal returns to effort, performance will decrease. Because we are unable to predict how employees will shift their effort from a marginal return to effort perspective, we present research questions. The research questions hold the motivation effect of RPI constant by making comparisons within each level of RPI, thereby allowing us to isolate how the effort distortion effect impacts performance. We formally state our research questions below:

RQ1a: When employees have private relative performance information, will performance differ when they cannot alter their effort allocations across tasks compared to when they can alter their effort allocations across tasks?

RQ1b: When employees have public relative performance information, will performance differ when they cannot alter their effort allocations across tasks compared to when they can alter their effort allocations across tasks?

III. METHOD

Experimental Design and Task Description

We use a 2×3 between-subjects experimental design in which we vary two factors. The first factor is whether participants can choose how much effort to allocate between a math task and a verbal problem-solving task that is manipulated at two levels: *No Choice* and *Choice*. The second factor is the RPI participants receive, that is manipulated at three levels: *No RPI*, *Private RPI*, and *Public RPI*. Each experimental session consists of four rounds (*Round* is a within-subjects factor) and five participants.

We use two primary dependent variables. *Total Problems Solved* is our measure of performance, which equals the number of problems solved correctly, regardless of task type. *Time Difference* is our measure of effort allocation distortion from the firm-preferred equal allocation, which is computed as the absolute value of the number of seconds spent on the math task minus the number of seconds spent on the verbal task. *Time Difference* is a continuous measure of the degree of effort distortion as opposed to a binary variable that simply captures whether individual participants ever distort their effort allocations.

Participants perform two problem-solving tasks, math and verbal, using an individual computer terminal.[3] The math task consists of solving multiplication problems and the verbal task consists of solving anagrams. Both types of problems must be solved without using any outside aids, including writing materials. Because the two tasks require different sets of abilities, a participant who performs worse than others on one task may still maintain a positive self-image by performing better than others on the other task.

In each round, participants are given 20 math problems and 20 verbal problems with five math problems and five verbal problems at each of four difficulty levels. Each math problem is classified at a specific difficulty level based on two criteria: (a) number of digits in the two multiplicands (e.g., multiplying a one-digit and a two-digit number versus multiplying two, two-digit numbers) and (b)

[3] The experiment was programmed using ZTree software (Fischbacher 2007).

frequency of carry-overs required to solve the problem. Each verbal problem is classified at a specific level of difficulty based on the length of the word (4, 5, 6, and 7 letters). To mitigate any effect of vocabulary differences across participants, the anagrams consist of common English words that should be readily recognizable once the letters have been correctly re-arranged.[4] We present the math and verbal problems in the same difficulty level order to all participants and explain this presentation order to the participants. Participants choose whether to start with the math or verbal problems, solve the problems in any order within a difficulty level and are free to move on to the next level without successfully completing all problems at a given level. They are prohibited from returning to a given problem type and difficulty level once they have moved on to a different type of problem or difficulty level.

The use of math and verbal problems at four difficulty levels creates a diminishing marginal return to effort for each specific task because problems that are more difficult are also more time consuming. In addition to capturing a key component of many real-world multi-task environments, differing difficulty levels also allows participants to draw stronger inferences about their own abilities, which, according to Festinger (1954), is the main reason to engage in social comparison.

Choice Manipulation

Participants in all conditions are allotted six minutes per round and, as described later, are paid on effort input such that they maximize compensation by allocating effort at three minutes per task. A clock is displayed on the computer screen so that participants can readily monitor the amount of time spent on each task. In the *No Choice* condition, the computer program displays each task for precisely three minutes each round, thereby enforcing the three-minute per task allocation. In the *Choice* condition, participants are given six minutes per round for both tasks and are free to allocate their time between tasks. The computer program displays a warning message when three minutes have expired to ensure that participants in the choice condition are aware of how they have allocated their time. However, they are not forced to switch to the other task.

RPI Manipulation

In all conditions, participants receive *individual* performance information at the end of each round showing how many math and verbal problems they solved correctly in the current round and in all rounds completed so far in the session. Participants in the *No RPI* condition receive no additional performance information. Participants in the *Private RPI* and *Public RPI* conditions receive cumulative performance rank information at the end of each round. The cumulative performance rank information is provided separately for the math and verbal tasks and is based on the relative performance of the five participants in each session. Participants in the *Private RPI* condition learn only their own cumulative ranks for the math and verbal tasks. Participants in the *Public RPI* condition learn both their own cumulative ranks and the cumulative ranks of each of the other four participants for each task. In both the *Private* and *Public RPI* conditions, participants with identical performance are given the same rank.

As described in the previous paragraph, participants in the *Private RPI* and *Public RPI* conditions learn their performance ranks for each task separately, but do not learn their overall performance rank for the two tasks combined. If we had also provided participants with their overall performance rank, we would be unable to disentangle whether shifts in effort were due to a desire to increase the performance rank on the emphasized task (as our theory predicts) or by a

[4] We pre-tested the multiplication and anagram problems to ensure that the problems assigned to each difficulty level took approximately the same amount of time to solve, on average, for the math task as for the verbal task.

desire to increase overall performance rank. Further, we provide RPI in the form of performance rank information rather than detailed performance distribution information because such detailed information could potentially foster the emergence of normative standards (Jackson and Harkins 1985; Paulus and Dzindolet 1993). The potential for such normative performance standards would make it difficult to determine if shifts in effort were motivated by a desire to increase performance rank on the emphasized task (as our theory predicts) or by a desire to conform to such standards.

Incentive System

Participants receive a show-up fee of $8.00. Additional compensation is based on effort allocation input rather than performance output. Specifically, participants are paid $0.013 for every second spent on each task, up to three minutes per task per round. Because participants in the *No Choice* condition must spend three minutes on each task per round, by design they earn $4.68 per round or $18.72 for the experiment (plus their $8.00 show-up fee). Participants in the *Choice* condition can spend more than three minutes on one of the tasks; however, doing so reduces their total compensation $0.013 per second for each second over three minutes spent on either task.

We do not base pay on performance output because the primary focus of our study is identifying the behavioral effects of RPI in multi-task environments, especially as these effects relate to the effort distortion effect. Including performance-output-based compensation in our experiment would create multiple incentives for participants to distort their effort as they consider both their utility for wealth and their utility for social distinction. This would weaken the theoretical contribution of our study because we would be unable to distinguish between the financial and behavioral motivations underlying any observed effort distortion.[5]

We provide a financial incentive to allocate effort equally across tasks in order to make these firm-preferred proportions salient and to counter-balance any intrinsic task interest that may motivate distorted effort allocations. Because this incentive imposes a cost for distorting effort allocations, any observed distortions can be clearly attributed to behavioral factors. The financial cost associated with allocating effort unequally also provides a more stringent test of our behavioral factors of interest than if no such cost were imposed.

Experimental Procedures

We conducted separate sessions for each of the six experimental cells. For each session, we recruited a cohort of participants from the same honors-based business course or business-based honor society at Georgia State University, a large public university in the United States. We recruited participants in cohorts from these courses and societies because it increased the likelihood that participants knew each other and would find the RPI to be more meaningful. We begin each session by providing participants with a copy of the instrument, which we read aloud to the participants. Participants then take a pre-experiment quiz and are required to score 100 percent on the quiz to ensure that they understand the experiment. Next, participants complete a pre-experiment questionnaire to measure self-esteem because prior research has shown that self-esteem may affect how individuals engage in and respond to social comparison (Gibbons and McCoy

[5] This design choice should not be interpreted as suggesting that behavioral effects would not exist in the presence of performance-based compensation. Behaviorally motivated effort distortions would still occur if the marginal utility for social distinction on a given task is greater than the marginal utility for wealth on an alternate task. The outcome of the trade-off in marginal utilities would depend on the specific parameters, i.e., the importance of peer status and the amount of incentive pay, of the setting. As discussed in the "Conclusions" section, addressing the trade-off between behavioral and financial factors is beyond the scope of our study.

1991).[6] Next, participants introduce themselves by standing and saying their participant number, which is prominently displayed on top of each participant's computer, his or her first name, and how long he or she has been a member of the honors' college or honor society. These introductions are made in order for participants in the *Public RPI* condition to know the identity of the participant associated with each performance rank. Introductions take place in all conditions to control for any potential effect of introductions not related to RPI. Participants then complete the math and verbal tasks in each of the four rounds of the experiment. Next, participants complete a post-experiment questionnaire. Finally, participants receive their earnings from the experiment via envelopes using participant identification numbers to ensure anonymity. Participants earned an average of $25.16, including the show-up fee, for approximately one hour of participation.

Participants

We conducted 18 sessions (three per experimental cell), with five participants in each session, for a total of 90 participants. The mean age of the participants is 24.3 years, and 53.3 percent are female. There are no significant age or gender differences across conditions (all $p > 0.12$, two-tailed). There is also no significant difference in pre-existing levels of self-esteem across conditions ($p = 0.18$, two-tailed). Therefore, we do not include age, gender, or self-esteem in our analyses.

IV. RESULTS

Measures and Descriptive Statistics

Table 1 reports descriptive statistics by experimental cell for the dependent variables used to test our hypotheses, *Total Problems Solved*, and *Time Difference*.[7] For completeness, we also report the number of *Problems Solved* and the *Time Spent* by task. Regarding *Time Difference*, an equal effort allocation results in a difference of zero and larger amounts represent greater distortions in effort allocations, with a maximum of 1440 if all time were spent on a single task.[8] For the *No Choice* condition, the mean *Time Difference* equals 0 by design.

As reported in Panel A of Table 1 and graphed in Figure 1, within the *No Choice* condition *Total Problems Solved* is directionally higher for *Private RPI* compared to *No RPI* (67.13 versus 57.00) as well as *Public RPI* compared to *Private RPI* (81.13 versus 67.13). This pattern of results is consistent with H1.

---

[6] To measure self-esteem, participants report how confident they are about their scholastic abilities; to what extent they feel that others respect and admire them; and how concerned they are about the impressions that they make. Participants respond on a Likert scale ranging from 1 to 7.

[7] Recall that *Total Problems Solved* is the summation of the correct output for the math and verbal tasks. This summation is an appropriate measure of performance because evidence indicates that the math and verbal problems are of the same difficulty level and, therefore, measure how effort translates into performance regardless of task. Within the *No Choice* condition, where the diminishing returns to effort should be the same across tasks, there is no significant difference in the mean time spent to solve math compared to verbal problems (22.47 seconds for math; 22.25 seconds for verbal; $t = 0.24$, $p = 0.81$, two-tailed). There are also no significant differences in the mean time spent per problem by task within the four difficulty levels (for math and verbal respectively, level 1: 16.51 and 18.33; level 2: 25.93 and 27.87; level 3: 36.21 and 36.43; level 4: 41.64 and 43.96; all p's > 0.12). In addition, post-experiment questionnaire responses indicate that participants did not perceive one task to be more difficult than the other ($t = 1.43$, $p = 0.16$, two-tailed).

[8] Participants were permitted to submit their answers before the round ended, but if they attempted to do so, a warning appeared on the screen that they would not be compensated for the remaining time of the current round. Because two participants in the *Private RPI* condition elected to submit their answers before the end of a round despite this warning, the summation of the mean time spent on the math and verbal tasks is less than 1440 in this condition. All inferences are the same if these two participants are excluded from our statistical analyses.

TABLE 1

Descriptive Statistics

Panel A: Math and Verbal Tasks Combined (Mean and Standard Deviation)

	No Choice			*Choice*		
	***No RPI* (n = 15)**	***Private RPI* (n = 15)**	***Public RPI* (n = 15)**	***No RPI* (n = 15)**	***Private RPI* (n = 15)**	***Public RPI* (n = 15)**
Total Problems Solved	57.00	67.13	81.13	54.73	70.00	71.00
	(8.15)	(11.39)	(14.27)	(13.70)	(16.02)	(19.97)
Time Difference	NA	NA	NA	64.86	202.38	625.23
				(61.13)	(231.80)	(484.03)

Panel B: Math Task (Mean, Standard Deviation, and Range)

	No Choice			*Choice*		
	***No RPI* (n = 15)**	***Private RPI* (n = 15)**	***Public RPI* (n = 15)**	***No RPI* (n = 15)**	***Private RPI* (n = 15)**	***Public RPI* (n = 15)**
Problems Solved	28.33	33.07	39.93	29.73	36.60	42.67
	(6.26)	(4.59)	(6.54)	(6.07)	(8.59)	(17.80)
	17–41	25–40	29–49	18–38	24–52	0–73
Time Spent	NA	NA	NA	729.82	709.06	839.63
				(43.93)	(164.22)	(384.64)
				633–821	391–1087	0–1440

Panel C: Verbal Task (Mean, Standard Deviation, and Range)

	No Choice			*Choice*		
	***No RPI* (n = 15)**	***Private RPI* (n = 15)**	***Public RPI* (n = 15)**	***No RPI* (n = 15)**	***Private RPI* (n = 15)**	***Public RPI* (n = 15)**
Problems Solved	28.67	34.07	41.20	25.00	33.40	28.33
	(3.48)	(10.53)	(10.83)	(9.37)	(14.21)	(14.44)
	24–35	19–63	26–70	11–45	7–60	0–57
Time Spent	NA	NA	NA	709.92	716.40	600.37
				(44.50)	(150.15)	(384.64)
				616–807	353–1049	0–1440

RPI (Relative Performance Information) is manipulated between-subjects at three levels: *No RPI*, *Private RPI*, and *Public RPI*. Participants in the *No RPI* condition do not receive any RPI. Participants in the *Private RPI* condition are provided with their own performance rank at the end of each round. Participants in the *Public RPI* condition are provided with their own performance rank and the performance ranks of the other four participants in the session at the end of each round.

Choice is manipulated between-subjects by whether participants are allowed to choose how to allocate their time between tasks. Participants in the *No Choice* condition are not allowed to choose how to allocate their time between tasks. Participants in the *Choice* condition are allowed to choose how to allocate their time between tasks.

Time Difference is the difference in the amount of time participants allocated between the math and verbal tasks, computed as the absolute value of the number of seconds spent on the math task minus the number of seconds spent on the verbal task. *Time Difference* equals 0 in the *No Choice* condition by design.

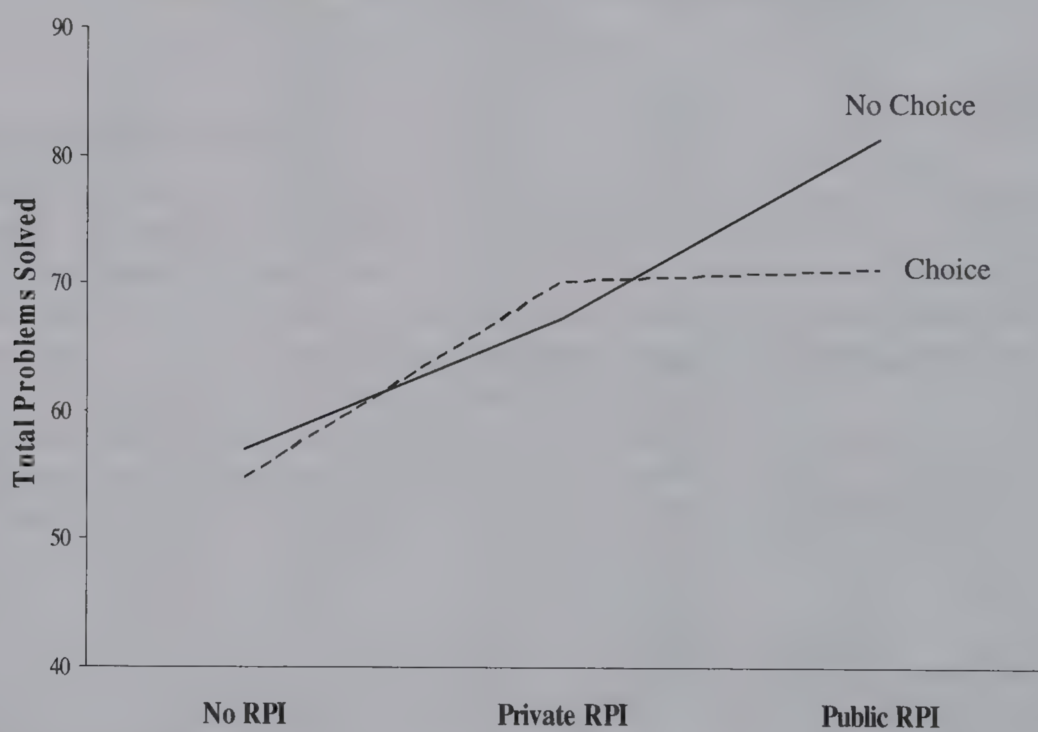
(continued on next page)

TABLE 1 (continued)

Total Problems Solved equals the total number of math and verbal problems solved in the four rounds of the experiment. *Problems Solved* equals the average number of math/verbal problems solved in the four rounds of the experiment. *Time Spent* equals the average number of seconds participants spent solving math/verbal problems in the four rounds of the experiment.

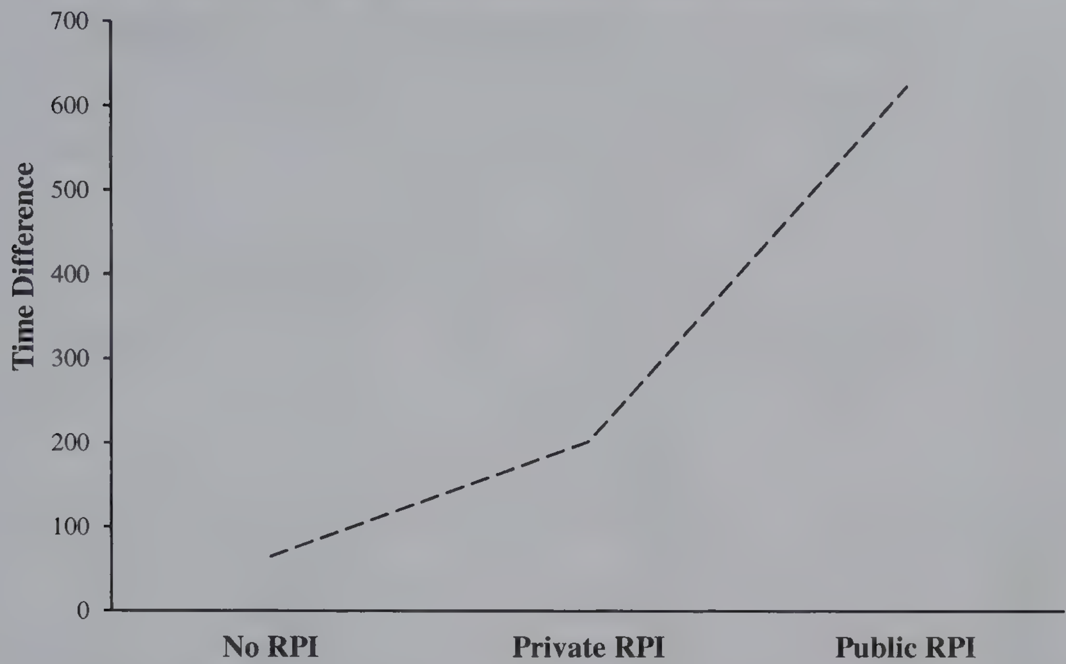
As reported in Panel A of Table 1 and graphed in Figure 2, within the *Choice* condition, *Time Difference* is directionally higher for *Public RPI* compared to *Private RPI* (625.23 versus 202.38) as well as *Private RPI* compared to *No RPI* (202.38 versus 64.86). This pattern of results is consistent with H2. To gain further insight into how individual participants distorted their effort allocations, Figure 3 presents histograms of *Time Difference* for each RPI condition. These histograms show that the distribution is highly skewed toward zero in the *No RPI* condition, is less skewed in the

FIGURE 1
Effect of RPI and Choice on Total Problems Solved



RPI (Relative Performance Information) is manipulated between-subjects at three levels: *No RPI*, *Private RPI*, and *Public RPI*. Participants in the *No RPI* condition do not receive any RPI. Participants in the *Private RPI* condition are provided with their own performance rank at the end of each round. Participants in the *Public RPI* condition are provided with their own performance rank and the performance ranks of the other four participants in the session at the end of each round.
Choice is manipulated between-subjects by whether participants are allowed to choose how to allocate their time between tasks. Participants in the *No Choice* condition are not allowed to choose how to allocate their time between tasks. Participants in the *Choice* condition are allowed to choose how to allocate their time between tasks.
Total Problems Solved equals the total number of math and verbal problems solved in the four rounds of the experiment.

FIGURE 2
Effect of RPI on Time Difference in the Choice Condition



RPI (Relative Performance Information) is manipulated between-subjects at three levels: *No RPI*, *Private RPI*, and *Public RPI*. Participants in the *No RPI* condition do not receive any RPI. Participants in the *Private RPI* condition are provided with their own performance rank at the end of each round. Participants in the *Public RPI* condition are provided with their own performance rank and the performance ranks of the other four participants in the session at the end of each round.
Choice is manipulated between-subjects by whether participants are allowed to choose how to allocate their time between tasks. Participants in the *No Choice* condition are not allowed to choose how to allocate their time between tasks. Participants in the *Choice* condition are allowed to choose how to allocate their time between tasks.
Time Difference is the difference in the amount of time participants allocated between the math and verbal tasks, computed as the absolute value of the number of seconds spent on the math task minus the number of seconds spent on the verbal task. *Time Difference* equals 0 in the *No Choice* condition by design.

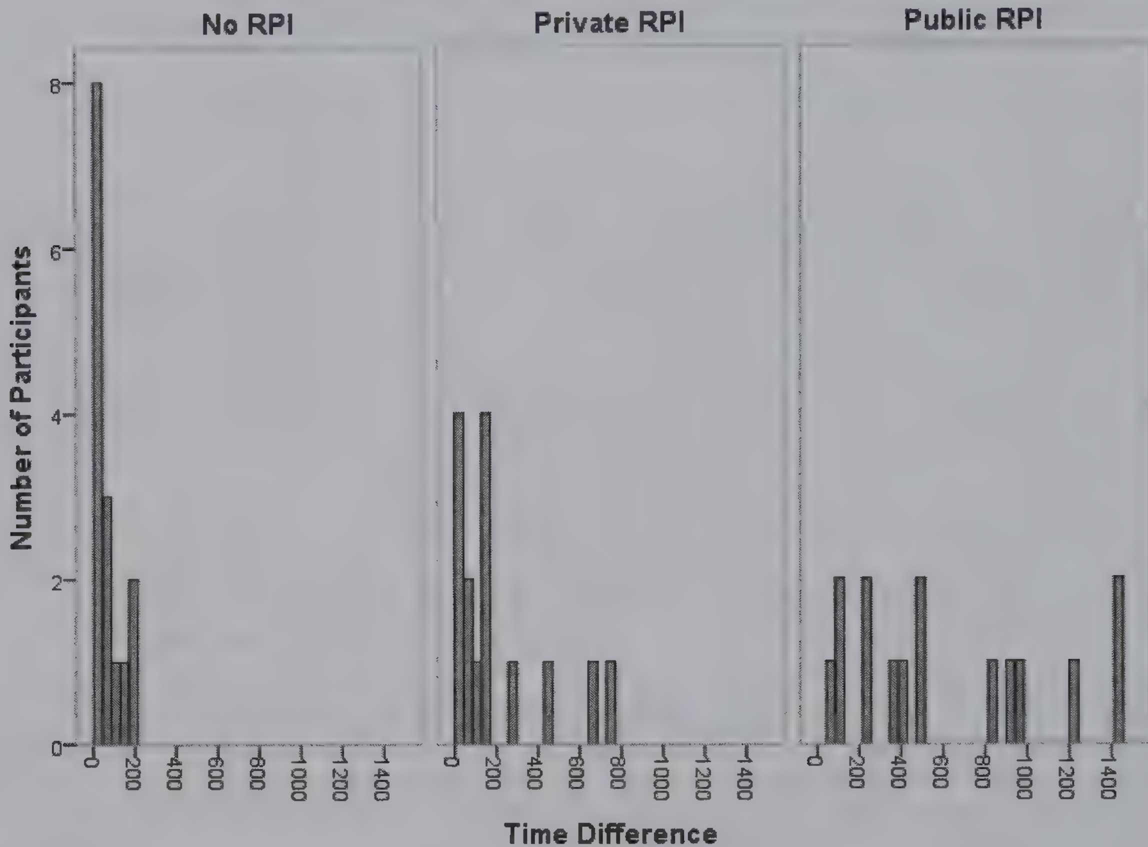
Private RPI condition, and is fairly uniform across the range in the *Public RPI* condition. This pattern of results is also consistent with H2.

Our research questions address the effect of choice on performance within the *Private* and *Public RPI* conditions. Panel A of Table 1 shows that *Total Problems Solved* is directionally greater for *Choice* compared to *No Choice* in the *Private RPI* condition (70.00 versus 67.13) but directionally lower for *Choice* compared to *No Choice* in the *Public RPI* condition (71.00 versus 81.13).

Tests of Hypotheses and Research Questions

To isolate the motivation effect of RPI, H1 considers the effect of RPI when effort allocations cannot be distorted across tasks. Therefore, our tests are performed within the *No Choice* condition. H1 predicts that employees with public RPI will perform better than those with private RPI and employees with private RPI will perform better than those without RPI, i.e., *Public RPI* > *Private RPI* > *No RPI*. We use planned contrasts (Buckless and Ravenscroft 1990) to formally test whether the results fall into the pattern predicted by H1. Specifically, we use contrast coefficients of +2 for

FIGURE 3
Time Difference Histograms for the Choice Condition



RPI (Relative Performance Information) is manipulated between-subjects at three levels: *No RPI*, *Private RPI*, and *Public RPI*. Participants in the *No RPI* condition do not receive any RPI. Participants in the *Private RPI* condition are provided with their own performance rank at the end of each round. Participants in the *Public RPI* condition are provided with their own performance rank and the performance ranks of the other four participants in the session at the end of each round.
Choice is manipulated between-subjects by whether participants are allowed to choose how to allocate their time between tasks. Participants in the *No Choice* condition are not allowed to choose how to allocate their time between tasks. Participants in the *Choice* condition are allowed to choose how to allocate their time between tasks.
Time Difference is the difference in the amount of time participants allocated between the math and verbal tasks, computed as the absolute value of the number of seconds spent on the math task minus the number of seconds spent on the verbal task. *Time Difference* equals 0 in the *No Choice* condition by design.

the *Public RPI* condition, +1 for the *Private RPI* condition, and −3 for the *No RPI* condition. The dependent variable is *Total Problems Solved*.

As reported in Table 2, Panel A, H1 is supported ($F = 27.46$, $p < 0.01$, one-tailed).[9] This result suggests that when employees do not have discretion over how to allocate their effort across tasks,

[9] Alternative sets of contrast coefficients for the predicted pattern yield identical inferences (all $p < 0.01$, one-tailed). Further, we perform an analysis of residual variance to test whether our planned contrast model is a good fit for the data. The residual sum of squares represents the between-group variance unexplained by the contrast model used to test H1. An insignificant p-value ($p = 0.23$) indicates that the contrast model explains substantially all of the between-group variance in the data.

TABLE 2

Test of Hypotheses

Panel A: H1: Dependent Variable—*Total Problems Solved*

Source of Variation	Df	MS	F-statistic	p-value[a]
Model Contrast[b]	1	372.50	27.46	$p < 0.01$
Error	42	13.57		

Panel B: H2: Dependent Variable—*Time Difference*

Source of Variation	Df	MS	F-statistic	p-value
Model Contrast[c]	1	429,810.28	17.47	$p < 0.01$
Error	42	24,602.76		

[a] p-values are one-tailed, given the directional predictions for these effects.
[b] Contrast coefficients are +2 for the *No Choice/Public RPI* condition, +1 for the *No Choice/Private RPI* condition, and −3 for the *No Choice/No RPI* condition.
[c] Contrast coefficients are +2 for the *Choice/Public RPI* condition, +1 for the *Choice/Private RPI* condition, and −3 for the *Choice/No RPI* condition.
RPI (Relative Performance Information) is manipulated between-subjects at three levels: *No RPI*, *Private RPI*, and *Public RPI*. Participants in the *No RPI* condition do not receive any RPI. Participants in the *Private RPI* condition are provided with their own performance rank at the end of each round. Participants in the *Public RPI* condition are provided with their own performance rank and the performance ranks of the other four participants in the session at the end of each round.
Choice is manipulated between-subjects by whether participants are allowed to choose how to allocate their time between tasks. Participants in the *No Choice* condition are not allowed to choose how to allocate their time between tasks. Participants in the *Choice* condition are allowed to choose how to allocate their time between tasks.
Total Problems Solved equals the total number of math and verbal problems solved in the four rounds of the experiment.
Time Difference is the difference in the amount of time participants allocated between the math and verbal tasks, computed as the absolute value of the number of seconds spent on the math task minus the number of seconds spent on the verbal task. *Time Difference* equals 0 in the *No Choice* condition by design.

knowing their own performance rank, i.e., private RPI, increases performance and the increase is even greater when employees know the performance rankings of all peers, i.e., public RPI.

H2 considers the effect of RPI on effort allocations within the *Choice* condition. H2 predicts that employees with public RPI will distort their effort allocations more than those with private RPI and employees with private RPI will distort their effort allocations more than those with no RPI, i.e., *Public RPI* > *Private RPI* > *No RPI*. Our test of H2 is similar to that of H1 in that we use planned contrasts with coefficients of +2, +1, and −3 for the *Public*, *Private*, and *No RPI* conditions, respectively, to test the pattern of results predicted by H2. The dependent variable is *Time Difference*.

As reported in Table 2, Panel B, H2 is supported ($F = 17.47$, $p < 0.01$, one-tailed).[10] We also note that *Time Difference* in the *No RPI* condition is not significantly different from zero ($t = 0.78$, $p = 0.44$, two-tailed), indicating that effort allocations across tasks were not distorted in the absence of RPI. These results suggest that RPI may induce employees to shift their effort toward particular tasks in order to reap the implicit benefits associated with receiving positive relative performance

[10] Alternative sets of contrast coefficients for the predicted pattern yield identical inferences (all $p < 0.01$, one-tailed). An insignificant p-value ($p = 0.14$, two-tailed) for the residual sum of squares indicates that the contrast model used to test H2 explains substantially all of the between-group variance in the data.

feedback and the shift may be more pronounced when RPI is public compared to private.[11,12] Participants in our study were willing to incur a financial cost in order to reap these implicit benefits. The mean financial cost incurred in the *No RPI*, *Private RPI*, and *Public RPI* conditions was \$1.29, \$2.10, and \$5.15, respectively.

Finally, we isolate the impact of the effort distortion effect on performance by comparing how choice affects performance within both the *Private* and *Public RPI* conditions. To establish a baseline, we first compare *Total Problems Solved* within the *No RPI* condition and find no statistically significant difference between the *No Choice* and *Choice* conditions (57.00 versus 54.73, $t = 0.87$, $p = 0.39$, two-tailed). This provides evidence that choice itself has no effect on performance, thereby allowing us to attribute any decrease in performance across choice conditions when RPI is present to the effort distortion effect of private or public RPI. Next, we isolate the impact of the effort distortion effect within the *Private RPI* condition (RQ1a). As reported in Table 3, *Total Problems Solved* is not significantly different between the *No Choice* and *Choice* conditions (67.13 versus 70.00, $t = 1.10$; $p = 0.28$, two-tailed). Thus, performance is not affected by effort distortions when RPI is private. Finally, we isolate the impact of the effort distortion effect within the *Public RPI* condition (RQ1b). As reported in Table 3, *Total Problems Solved* is greater in the *No Choice* condition than in the *Choice* condition (81.13 versus 71.00, $t = 3.88$; $p < 0.01$, two-tailed).[13] This overall pattern of results shows that the effort distortion effect of RPI leads to decreased performance in our experiment only when RPI is public.

In summary, the results of our experiment suggest that in a multi-task environment, RPI has both motivation and effort distortion effects. The motivation effect is isolated when participants have no discretion over how to allocate their effort across tasks. In the *No Choice* condition we find that both private and public RPI increases performance, with public RPI resulting in a greater performance increase. When participants have discretion in determining effort allocations, however, both the motivation and effort distortion effects play a role in determining performance. Both private and public RPI result in distorted effort allocations across tasks, and public RPI exacerbates this distortion. In our setting, where there are diminishing marginal returns to effort, the exacerbated distortion effect in the *Public RPI* condition leads to decreased performance.

---

[11] To test whether participants in the *Choice* condition were more likely to distort effort allocations toward one of the tasks in particular, we compare the mean time spent on math and verbal problems by individual participants. Results of a paired-samples t-test show that these times are not significantly different ($t = 1.16$, $p = 0.25$, two-tailed), indicating that no task was preferred overall. We also conduct the test within each RPI condition and confirm that participants, on average, did not spend significantly more time on math or verbal tasks in any of the RPI conditions (all $p > 0.24$, two-tailed).

[12] A potential alternative explanation for the effort distortion effect is that participants allocated more effort to the task they found more interesting. We reject this alternative explanation for two reasons. First, there is no difference in effort allocations in the absence of RPI. Second, responses from the post-experiment questionnaire indicate that effort allocations are not affected by perceptions of task interest. Specifically, we asked participants to assess how interesting they found the math and verbal tasks on a scale ranging from 1 (not at all interesting) to 7 (very interesting). A regression using the difference between the time spent on the math and the verbal tasks as the dependent measure and the difference between perceived math and verbal task interest as the independent measure shows that participants' perceptions of relative task interest do not affect their effort allocations ($t = 0.12$, $p = 0.90$, two-tailed).

[13] On the surface, it may seem counter-intuitive that allowing participants to distort their effort toward a task for which they presumably believe they can perform better resulted in fewer total problems solved than when participants were unable to choose how to allocate their effort. This result can be explained by a combination of the diminishing marginal return to effort present in each of our tasks and participants' preference for self-affirmation and positive social distinction. That is, the difficulty of a given problem type increases as more problems of that type are attempted. As such, returns to effort on a given task decrease the more problems of that type that are attempted because each additional problem takes longer to solve. Participants' desire to distinguish themselves on at least one task, however, keeps them working on that task even as the returns to effort diminish rather than switching to the other task where the returns to effort would be greater, at least initially.

TABLE 3

Tests of Research Questions
Dependent Variable—*Total Problems Solved*

Research Question	*Total Problems Solved* (Standard Deviation) *No Choice*	*Choice*	t-statistic	p-value[a]
RQ1a: *Private RPI*	67.13	70.00	1.10	0.28
	(11.39)	(16.02)		
RQ1b: *Public RPI*	81.13	71.00	3.88	< 0.01
	(14.27)	(19.97)		

[a] p-values are two-tailed, given the lack of directional predictions for these effects.
RPI (Relative Performance Information) is manipulated between-subjects at three levels: *No RPI*, *Private RPI*, and *Public RPI*. Participants in the *No RPI* condition do not receive any RPI. Participants in the *Private RPI* condition are provided with their own performance rank at the end of each round. Participants in the *Public RPI* condition are provided with their own performance rank and the performance ranks of the other four participants in the session at the end of each round.
Choice is manipulated between-subjects by whether participants are allowed to choose how to allocate their time between tasks. Participants in the *No Choice* condition are not allowed to choose how to allocate their time between tasks. Participants in the *Choice* condition are allowed to choose how to allocate their time between tasks.
Total Problems Solved equals the total number of math and verbal problems solved in the four rounds of the experiment.

Supplemental Analysis

Post-Experiment Questionnaire Data

We use post-experiment questionnaire data in conjunction with our observed dependent variables to provide support for the theory underlying our results. All analyses in this section pool the *Choice* and *No Choice* conditions.

First, we investigate whether private RPI motivates participants to be concerned about their performance relative to their peers and whether public RPI magnifies this concern. We use three questions in the post-experiment questionnaire to capture participants' concern for social comparison and report the mean response to each question by experimental cell in Table 4.[14] We create a social comparison factor using the three questions (eigenvalue = 1.96, percentage of variance explained = 65.46) and then conduct a planned contrast test with the social comparison factor as the dependent variable and the three levels of RPI as the independent variable. Consistent with the predicted pattern and the planned contrasts used for testing H1 and H2, the contrast coefficients are +2 for the *Public RPI* condition, +1 for the *Private RPI* condition, and −3 for the *No RPI* condition. Untabulated results show that the effect of RPI on social comparison is in the predicted pattern ($t = 2.72$, $p < 0.02$, two-tailed). These results are consistent with our theory that private RPI, and especially public RPI, motivates employees to be concerned about how well they perform relative to their peers.

[14] Participants reported how often they thought about how their performance ranked relative to those of the other participants in the experiment (*Rank Thinking*); to what extent they were nervous or concerned about how well they were performing relative to other participants in the experiment (*Rank Nervousness*); and to what extent thinking about performance comparison interfered with their ability to concentrate on the problems (*Rank Interference*). Participants responded on a scale ranging from 1 (Not at all) to 7 (To a great extent).

TABLE 4

Responses to Post-Experiment Questions (Mean and Standard Deviation)

	No Choice			*Choice*		
	No RPI	*Private RPI*	*Public RPI*	*No RPI*	*Private RPI*	*Public RPI*
Rank Thinking	3.80	4.93	5.40	3.93	4.53	5.27
	(1.57)	(1.44)	(1.18)	(1.94)	(1.96)	(1.10)
Rank Nervousness	3.67	3.93	5.00	3.40	3.40	4.40
	(1.72)	(1.10)	(1.69)	(1.64)	(1.99)	(1.45)
Rank Interference	2.33	2.73	3.73	2.87	2.33	2.73
	(1.68)	(1.62)	(2.15)	(1.77)	(1.23)	(1.91)
Pride	4.00	4.20	4.40	4.27	4.47	4.07
	(0.85)	(1.37)	(1.84)	(1.16)	(1.69)	(1.87)
Feelings Strength	0.40	1.13	1.60	0.66	1.40	1.67
	(0.74)	(0.74)	(0.91)	(0.97)	(0.99)	(0.73)

RPI (Relative Performance Information) is manipulated between-subjects at three levels: *No RPI*, *Private RPI*, and *Public RPI*. Participants in the *No RPI* condition do not receive any RPI. Participants in the *Private RPI* condition are provided with their own performance rank at the end of each round. Participants in the *Public RPI* condition are provided with their own performance rank and the performance ranks of the other four participants in the session at the end of each round.

Choice is manipulated between-subjects by whether participants are allowed to choose how to allocate their time between tasks. Participants in the *No Choice* condition are not allowed to choose how to allocate their time between tasks. Participants in the *Choice* condition are allowed to choose how to allocate their time between tasks.

Rank Thinking represents participants' assessments of the extent to which they thought about performance rank relative to other participants in the study, using a seven-point Likert scale (1 = never, 4 = sometimes, 7 = very often).

Rank Nervousness represents participants' assessments of how nervous or concerned they were about how well they were performing relative to the other participants, using a seven-point Likert scale (1 = not at all nervous or concerned, 4 = somewhat nervous or concerned, 7 = very nervous or concerned).

Rank Interference represents participants' assessments of how their performance rank compared to other participants' rank interfered with their ability to concentrate on the problems using a seven-point Likert scale (1 = not at all, 4 = somewhat, 7 = very much).

Pride represents participants' assessments of how they felt about their performance, using a seven-point Likert scale (1 = very ashamed, 4 = neither ashamed nor proud, 7 = very proud) and 4 (neither ashamed nor proud).

Feelings Strength represents the degree of shame or pride participants reported, computed as the absolute value of the difference between the midpoint of the *Pride* question, 4.0, and the participant's response. This computation results in a range of 0 to 3, with 3 representing the highest degree of shame or pride.

Second, we investigate whether high (low) relative performance creates feelings of pride (shame) and whether public RPI magnifies these feelings. We use one question on the post-experiment questionnaire to measure feelings of pride/shame and create a continuous variable that we refer to as *Pride*. The pride question measures how participants felt about their own performance on a scale ranging from 1 (very ashamed) to 7 (very proud). Mean *Pride* by cell is reported in Table 4. We compute a measure of relative performance by summing each participant's final math and verbal ranks, which we refer to as *Total Rank*. The *smaller* the total rank, the *better* the relative total performance. As such, we expect *Pride* and *Total Rank* to be negatively related when private or public RPI is present. We do not expect such a relation when RPI is absent because, although participants are aware of their absolute performance, they have no benchmark for evaluating their relative performance. Our analysis supports this predicted pattern. Specifically, *Pride* and *Total Rank* are negatively correlated in the *Private RPI* ($r = -0.81$, $p < 0.01$) and *Public RPI* ($r = -0.82$,

$p < 0.01$) conditions and there is no significant correlation in the *No RPI* condition ($r = -0.27$, $p = 0.15$).[15]

To investigate whether these feelings are magnified when participants receive public RPI, we compare the strength of their reported feelings of pride or shame, *Feelings Strength*, across the *Private RPI* and *Public RPI* conditions. *Feelings Strength* is computed as the absolute value of the difference between each participant's answer to the pride question and the neutral point of 4 (neither ashamed nor proud), thereby capturing extreme feelings of pride or shame. Mean *Feelings Strength* by cell is reported in Table 4. Results of a t-test show that *Feelings Strength* is higher in the *Public RPI* compared to the *Private RPI* condition ($t = 1.69$, $p < 0.05$, one-tailed).

The correlation and t-test results provide evidence that participants responded to RPI by feeling more proud (ashamed) when the information content indicated that their relative performance was better (lower) than their peers and that these feelings were magnified when the RPI was public. Because *Pride* was measured at the end of the experiment, we cannot unambiguously infer that participants' behavior during the experiment was affected by their anticipation of how they would feel about their relative performance. Nonetheless, these results are consistent with a key construct (i.e., that relative performance affects how one feels about oneself) underlying our theoretical process.

Finally, we investigate whether participants behave as if they have a preference for social distinction. Evidence of such a preference would take the form of participants feeling prouder the better they perform on one task, e.g., pride is greater for ranking first and third on the two tasks compared to ranking second on both tasks. We create a variable, *Rank Difference*, computed as the absolute value of each participant's math rank minus verbal rank. We expect that *Pride* and *Rank Difference* are positively related in the *Private RPI* and *Public RPI* conditions, but do not expect any relation in the *No RPI* condition because participants are unaware of their rankings. We use partial correlations for our analysis to control for *Total Rank*. As expected, this analysis shows a significant positive correlation between *Pride* and *Rank Difference* in both the *Private RPI* ($r = 0.49$, $p < 0.01$) and the *Public RPI* ($r = 0.35$, $p < 0.01$) conditions, but no significant correlation in the *No RPI* condition ($r = 0.24$, $p = 0.20$).[16] As with the previous analysis, we cannot unambiguously infer that participants' behavior during the experiment was affected by a preference for social distinction. However, this analysis provides evidence of a preference for social distinction and this preference is consistent with the causal link between RPI and distorted effort allocations.

First Round Data

The tests of our hypotheses use data pooled across the four rounds of the experiment. To investigate whether the effects of RPI are observed only after participants receive performance feedback or also exist initially (due to participants' ability to anticipate their likely RPI feedback), we conduct tests of all hypotheses and research questions using only first round data. We find that

---

[15] An analysis using Fisher's z-transformation shows that the correlation coefficient for the *No RPI* condition is significantly different from those of the *Private RPI* ($z = 3.26$, $p < 0.01$, two-tailed) and the *Public RPI* ($z = 3.37$, $p < 0.01$, two-tailed) conditions. For completeness, we also compare the *Private RPI* and *Public RPI* correlation coefficients and find that they are not significantly different ($p = 0.47$). This is not surprising because correlation coefficients are not influenced by the magnitude of the variables (DeGroot and Schervish 2002). In other words, even though we expect that feelings of pride and shame are magnified when RPI is public rather than private, these magnified feelings should not lead to a greater correlation coefficient between *Pride* and *Total Rank* in the *Public RPI* condition. Covariance, however, is influenced by the magnitudes of the variables (DeGroot and Schervish 2002) and should be greater in the *Public RPI* condition. Our analysis confirms this expectation ($p < 0.05$, two-tailed).

[16] Further analysis, using Fisher's z-transformation, shows that the correlation coefficient for the *No RPI* condition is significantly different from the correlation for the *Private RPI* condition ($z = 2.35$, $p < 0.03$, two-tailed) and significantly different at the marginal level from the correlation for the *Public RPI* condition ($z = 1.86$, $p < 0.03$, two-tailed).

all tests of hypotheses and research questions are replicated using only the first round data (all p's $<$ 0.02, for all hypotheses and RQ1b except RQ1a where p = 0.56).[17] These results imply that participants were able to anticipate their likely relative performance and responded accordingly.

Responses to First Round RPI

Results from the first round reveal that participants anticipated and responded to RPI before it was actually provided. To investigate whether participants in the *Choice* condition adjusted their effort allocations after actually receiving RPI, we create two new variables: *Initial Rank Difference* and *Change In Time Allocated. Initial Rank Difference* is computed as each participant's math rank minus verbal rank at the end of round 1. *Change in Time Allocated* is computed as the change in the percentage of time allocated to the verbal task between round 1 and the remaining three rounds. A regression, using *Change in Time Allocated* as the dependent variable and *Initial Rank Difference* as the independent variable, finds the effect of *Initial Rank Difference* is positive in both the *Private RPI* ($t = 1.79$, $p = 0.08$, two-tailed) and *Public RPI* ($t = 2.67$, $p < 0.02$, two-tailed) conditions, indicating that participants shifted their effort allocations toward the task where their relative performance was better in round 1.[18] There is no effect in the *No RPI* condition ($t = 1.31$, $p = 0.21$, two-tailed), which is expected given that participants were not provided with rank information. These results reveal that participants responded to RPI feedback in the first round by allocating additional time in subsequent rounds to the task where their own relative performance was better.

We also investigate how participants in the *Choice* condition responded to the content of the RPI they received at the end of the first round. We restrict our analysis to participants who did well on both tasks (rank 1 or 2 on both tasks) or poorly on both tasks (rank 4 or 5 on both tasks) to explore whether such content affected their effort allocations in subsequent rounds. This analysis shows that participants who received RPI indicating that they performed poorly on both tasks in round 1 shifted their focus from one task to the other in 50 percent of subsequent rounds, on average. In contrast, participants who received RPI indicating that they performed well on both tasks in round 1, shifted their focus in only 14 percent of subsequent rounds, on average. This difference is statistically significant ($t = 2.36$, $p < 0.04$, two-tailed). These results indicate that participants who did poorly on both tasks were less able to settle on an effort allocation strategy than participants who did well on both tasks.[19]

V. CONCLUSIONS

This paper reports the results of an experiment that investigates the behavioral effects of RPI on performance and the allocation of effort in a multi-task environment. We isolate the behavioral effects of RPI by using a setting in which employees are not compensated based on performance output but are compensated based on their effort allocations across tasks. Overall, the results of our experiment suggest that, in a multi-task environment, RPI has both motivation and effort distortion effects. We isolate the motivation effect by observing behavior when participants have no discretion over how to allocate their effort across tasks. The motivation effect of RPI results in higher

---

17 We also test all hypotheses and research questions without round 1, i.e., by using only round 2, 3, and 4 data. Results are inferentially identical to the tests reported with round 1 included.

18 Recall that higher ranks indicate lower relative performance. Therefore, if the difference between the math and verbal ranks in round 1 (*Initial Rank Difference*) is positive, then the participant ranked better on the verbal task. A shift toward the verbal task in subsequent rounds, therefore, indicates a shift toward the task where the participant's relative performance was better in the first round.

19 A natural question is whether this result differs when RPI is provided privately versus publicly. We are unable to answer this question due to low statistical power given that only 11 participants did well or poorly on both tasks in round 1.

performance for both private and public RPI, and, consistent with Tafkov (2011), the effect on performance is magnified when the RPI is public. When participants have discretion in determining effort allocations, however, both the motivation and effort distortion effects play a role in determining performance. We find that both private and public RPI results in distorted effort allocations, and public RPI exacerbates this distortion. Further, we demonstrate that the effort distortion effect may hurt total performance.

Our study provides insights on determining the costs and benefits of organizational decisions, which is the fundamental role of accounting (Balakrishnan et al. 2009), and has important implications for practice and theory. For example, our study suggests that if the firm is able to control how employees allocate their effort across tasks, then providing RPI is likely to increase performance because, similar to a single-task environment, only the motivation effect is present. However, when employees have discretion over how they allocate their effort, the RPI-induced effort distortion effect may make it more difficult to achieve the appropriate effort allocations, which in turn may affect performance. Specifically, we demonstrate that, in a setting where all tasks exhibit diminishing marginal returns to effort, the effort distortion effect caused by public RPI may have a negative effect on performance. However, in settings where the diminishing marginal returns to effort are less severe or even increasing, it is plausible that the RPI-induced effort distortion effect can complement the motivation effect, resulting in an increase rather than a decrease in performance. Thus, our study implies that firms need to consider both the nature of the tasks as well as how effort distortions may ultimately affect performance when making decisions regarding RPI. Our study also suggests that firms need to consider how to assign task responsibilities given the availability of informal RPI. Under such circumstances, firms may benefit from aggregating into single jobs only those tasks that require similar skills so as to mitigate any effort distortion effects.

Our study also contributes to two streams of research. Prior research that investigates how RPI affects performance has documented the positive motivating effects of RPI in the single-task environment (Kerr et al. 2007; Hannan et al. 2008; Tafkov 2011). Our study extends this stream by demonstrating that the results of this research cannot be simply generalized to the multi-task environment, where the overall effect of RPI on performance depends on the interplay between the motivation and effort distortion effects. We also contribute to the stream of research investigating the design of incentive contracts in multi-task environments (Holmstrom and Milgrom 1991; Feltham and Xie 1994; Datar et al. 2001; Farrell et al. 2008; Kachelmeier et al. 2008; Kachelmeier and Williamson 2010) as well as how social preferences such as fairness (Fehr and Schmidt 2004) or social pressure (Bruggen and Moers 2007) may influence the allocation of effort. Our study adds to this growing knowledge base by showing that RPI has an important influence as well.

Finally, our study suggests several avenues for future research. First, our focus in this study is on identifying the behavioral effects of RPI on employee motivation and effort allocations. As such, our setting excluded incentive pay based on performance output. In settings where employees are compensated based on performance output, however, their effort allocations would be affected by the financial returns to effort expended on a task in addition to the behavioral effects identified in this study. Future research could investigate how the interplay between financial incentives and behavioral factors affects effort allocations in multi-task environments. Second, so that we could clearly attribute effort allocation distortions to the participants' desire to increase their performance rank on the emphasized task, we provide participants in the *Private RPI* and *Public RPI* conditions rank information on each task separately but not overall rank information on the two tasks combined. This limits our ability to generalize our results to settings where employees receive overall rank information because effort allocations would affect not only their ranking on individual tasks, but also their overall rankings as well. Future research could investigate how providing overall rank information affects effort allocations. Third, participants received performance rank information rather than detailed performance distribution information. Since some firms provide

detailed performance information, it would be interesting to investigate how such information would affect both motivation and effort allocations. Finally, the tasks in our experiment have a diminishing marginal return to effort. Future research could investigate how performance and effort allocation are affected in settings where some or all tasks have increasing marginal returns to effort.

REFERENCES

Anderson, D., C. Crowell, S. Sponsel, M. Clarke, and J. Brence. 1982. Behavior management in the public accommodations industry: A tree-project demonstration. *Journal of Organizational Behavioral Management* 4: 33–66.

Askin, O. Z., and P. T. Harker. 1999. To sell or not to sell: Determining the trade-offs between service and sales in retail banking phone centers. *Journal of Service Research* 2 (1): 19–33.

Baker, G., M. Jensen, and K. Murphy. 1988. Compensation and incentives: Practice vs. theory. *Journal of Finance* 43 (3): 593–616.

Balakrishnan, R., K. Sivaramakrishnan, and G. Sprinkle. 2009. *Managerial Accounting*. New York, NY: John Wiley & Sons, Inc.

Beach, S., and A. Tesser. 1995. Self-esteem and the extended self-evaluation maintenance model: The self in social context. In *Efficacy, Agency, and Self-Esteem*, edited by M. Kernis, 145–170. New York, NY: Plenum Press.

Berger, L., K. Klassen, T. Libby, and A. Webb. 2011. *Evidence on the Relation between Incentives and Performance in Repeated Tournaments. Working paper*, University of Waterloo.

Brickman, P., and R. Bulman. 1977. Pleasure and pain in social comparison. In *Social Comparison Processes: Theoretical and Empirical Perspectives*, edited by J. Suls, and R. Miller, 149–186. Washington DC: Hemisphere.

Bruggen, A., and F. Moers. 2007. The role of financial incentives and social incentives in multi-task Settings. *Journal of Management Accounting Research* 19: 25–50.

Buckless, F., and S. Ravenscroft. 1990. Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review* 65: 933–945.

Datar, S., S. C. Kulp, and R. A. Lambert. 2001. Balancing performance measures. *Journal of Accounting Research* 39: 75–92.

DeGroot, M., and M. Schervish. 2002. *Probability and Statistics*. 3rd ed. Boston, MA: Addison-Wesley.

Dumont, E., B. Fortin, N. Jacquemet, and B. Shearer. 2007. *Physicians' Multitasking and Incentives: Empirical Evidence from a Natural Experiment*. Discussion Paper No. 3229. Bonn, Germany: Institute for Study of Labor.

Farrell, A. M., K. Kadous, and K. L. Towry. 2008. Contracting on contemporaneous versus forward-looking measures: An experimental investigation. *Contemporary Accounting Research* 25 (3): 773–802.

Fehr, E., and K. M. Schmidt. 2004. Fairness and incentives in a multi-task principal-agent model. *Scandinavian Journal of Economics* 106: 453–474.

Feltham, G. A., and J. Xie. 1994. Performance measure congruity and diversity in multi-task principal/agent relations. *The Accounting Review* 69: 429–453.

Festinger, L. 1954. A theory of social comparison. *Human Relations* 7: 117–140.

Fischbacher, U. 2007. ZTree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10: 171–178.

Frederickson, J. 1992. Relative performance information: The effect of common uncertainty and contract type on agent effort. *The Accounting Review* 67: 647–669.

Frey, B. S. 2007. Awards as compensation. *European Management Review* 4: 6–14.

Gibbons, F., and S. McCoy. 1991. Self-esteem, similarity and reactions to active versus passive downward comparison. *Journal of Personality and Social Psychology* 60: 414–424.

Goffman, E. 1959. *The Presentation of Self in Everyday Life*. New York, NY: Doubleday.

Hannan, R. L., R. Krishnan, and A. H. Newman. 2008. The effects of disseminating relative performance feedback in tournament versus individual performance compensation plans. *The Accounting Review* 83: 893–913.

Hoffman, P., L. Festinger, and D. Lawrence. 1954. Tendencies toward group comparability in competitive bargaining. *Human Relations* 7: 141–159.

Holmstrom, B., and P. Milgrom. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *The Journal of Law, Economics, and Organization* 7: 24–52.

Jackson, J., and S. Harkins. 1985. Equity in effort: An explanation of the social loafing effect. *Journal of Personality and Social Psychology* 49: 1199–1206.

Kachelmeier, S. J., B. Reichert, and M. G. Williamson. 2008. Measuring and motivating quantity, creativity, or both. *Journal of Accounting Research* 46 (2): 341–373.

Kachelmeier, S. J., and M. G. Williamson. 2010. Attracting creativity: The initial and aggregate effects of contract selection on creativity-weighted productivity. *The Accounting Review* 85 (5): 1669–1691.

Kandel, E., and E. Lazear. 1992. Peer pressure and partnership. *Journal of Political Economy* 100: 801–817.

Kerr, N., L. Messe, D. Seok, E. Sambolec, R. Lount, and E. Park. 2007. Psychological mechanisms underlying the Kohler motivation gain. *Personality and Social Psychology Bulletin* 33: 828–841.

Lazarus, R. 1991. *Emotion and Adaptation*. New York, NY: Oxford University Press.

Lindbeck, A., and D. J. Snower. 2000. Multitask learning and the reorganization of work: From Tayloristic to holistic organization. *Journal of Labor Economics* 18 (3): 353–376.

Lynn, M., and C. Snyder. 2002. Uniqueness seeking. In *Handbook of Positive Psychology*, edited by C. Snyder, and S. Lopez, 395–410. New York, NY: Oxford University Press.

Newman, A., and I. Tafkov. 2011. *Relative Performance Information in Tournaments with Different Prize Structures*. Working paper, University of Pittsburgh and Georgia State University.

Nordstrom, R., P. Lorenzi, and R. Hall. 1990. A Review of public posting of performance feedback. *Journal of Organizational Behavior Management* 11: 101–123.

O'Connell, V. O. 2008. Retailers reprogram workers in efficiency push. *Wall Street Journal* (September 10): B1.

Paulus, P., and M. Dzindolet. 1993. Social influence process in group brainstorming. *Journal of Personality and Social Psychology* 64: 575–586.

Silverman, R. E. 2011. Latest game theory mixes work and play. *Wall Street Journal* (October 10): B11.

Simsek, O., and B. Yalincetin. 2010. I feel unique, therefore I am: The development and preliminary validation of the personal sense of uniqueness (PSU) scale. *Personality and Individual Differences* 49: 576–581.

Smith, R. 2000. Assimilative and contrastive emotional reactions to upward and downward social comparisons. In *Handbook of Social Comparison: Theory and Research*, edited by J. Suls, and L. Wheeler, 173–200. New York, NY: Kluwer Academic/Plenum Publishers.

Smith, R., and C. Insko. 1987. Social comparison choices during ability evaluation: The effects of comparison publicity, performance feedback, and self-esteem. *Personality and Social Psychology Bulletin* 13: 111–122.

Snyder, C., and H. Fromkin. 1980. *Uniqueness: The Human Pursuit of Difference*. New York, NY: Plenum.

Steele, C. 1988. The psychology of self-affirmation: Sustaining the integrity of the self. In *Advances in Experimental Social Psychology*, edited by L. Berkowitz, 261–302. San Diego, CA: Academic Press.

Tadeschi, J. 1981. *Impression Management Theory and Social Psychological Research*. New York, NY: Academic Press.

Tafkov, I. 2011. *Private and Public Relative Performance Information under Different Incentive Systems*. Working paper, Georgia State University.

Tesser, A. 1988. Toward a self-evaluation maintenance model of social behavior. In *Advances in Experimental Social Psychology*, edited by L. Berkowitz, 181–228. San Diego, CA: Academic Press.

Wikoff, M., D. Anderson, and C. Crowell. 1982. Behavior management in a factory setting: Increasing work efficiency. *Journal of Organizational Behavior Management* 4: 97–127.

Wilson, S., and L. Benner. 1971. The effects of self-esteem and situation upon comparison choices during ability evaluation. *Sociometry* 34: 381–397.

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 577–609

American Accounting Association
DOI: 10.2308/accr-50338

Investor Perceptions of Potential IFRS Adoption in the United States

Philip P. M. Joos
Edith Leung
Tilburg University

ABSTRACT: This paper examines the stock market reaction to 15 events relating to IFRS adoption in the United States. The goal is to assess whether investors perceive the switch to IFRS as beneficial or costly. Our findings suggest that investors' reaction to IFRS adoption is more positive in cases where IFRS is expected to lead to convergence benefits. Our results also indicate a less positive market reaction for firms with higher litigation risk, which is consistent with investors' concerns about greater discretion and less implementation guidance under IFRS for these firms. Overall, the findings are relevant to the current debate on IFRS adoption in the U.S. and highlight the importance of convergence to investors.

Keywords: *IFRS adoption; convergence; SEC.*

Data Availability: *All data are publicly available from the sources indicated in the paper (see Appendices A and B).*

I. INTRODUCTION

This study investigates to what degree the U.S. stock market reacted to public events associated with the adoption of International Financial Reporting Standards (IFRS) by domestic U.S. firms. On April 24, 2007, the Securities and Exchange Commission (SEC) announced it was contemplating the mandatory use of IFRS by U.S. companies. The SEC's motivation was that U.S. investors would benefit from a single set of high-quality global standards. Although several studies have documented positive effects of IFRS adoption in Europe (Barth et al. 2008; Daske et al. 2008; Armstrong et al. 2010), it is unclear whether a switch to IFRS would be beneficial in the U.S. Since current U.S. accounting standards (i.e., U.S. GAAP) and U.S. reporting are generally considered to be of high quality (Leuz et al. 2003; Bradshaw et al. 2004), the switch

We are grateful to John Harry Evans III (senior editor), Wayne Thomas (editor), and two anonymous referees for their constructive feedback and detailed suggestions. We also appreciate helpful comments from Wayne Landsman, Christian Leuz, Peter Pope, Holly Skaife, and seminar participants at the 2010 AAA IAS Midyear Conference, 2010 EAA Doctoral Colloquium and Annual Conference, the HMMT Workshop at Maastricht University, Humboldt University, and Tilburg University.

Editor's note: Accepted by John Harry Evans III, with thanks to Wayne Thomas for serving as editor on previous versions.

Submitted: March 2010
Accepted: October 2012
Published Online: October 2012

may not provide significant benefits in terms of "higher quality" financial reporting. It is also unclear whether investors expect the switch to lead to convergence benefits, such as reduced costs of comparing firms' financial reporting globally (SEC 2008; Armstrong et al. 2010; Hail et al. 2010) and greater consistency of financial information by enabling auditors and their clients to develop consistent global practices to deal with accounting issues (Tweedie 2006), especially since U.S. GAAP and IFRS have become increasingly similar in recent years.

This study provides empirical evidence on how investors evaluate the potential switch to IFRS. We examine U.S. stock market reactions to events that affect the likelihood of IFRS adoption, similar to Christensen et al. (2007) and Armstrong et al. (2010).[1] If investors perceive IFRS adoption to be beneficial, we expect to observe a positive (negative) market reaction to events that increase (decrease) the likelihood of adoption. Our main analysis focuses on a differential effect of IFRS adoption across U.S. firms for which we make three predictions. First, we expect a lower market reaction if investors believe IFRS will adversely affect reporting quality due to the lack of implementation guidance, which is a particular concern for firms in the extractive and insurance industries and firms with high litigation risk.[2] Second, we predict that investors will react more positively if they expect IFRS to result in convergence benefits, which is more likely in industries where IFRS is already widely adopted by non-U.S. peer firms. Third, we expect investors' reaction to IFRS adoption to vary with the direct cost impact of introducing these standards and focus on firms that currently report under both U.S. GAAP and IFRS and those that apply LIFO.

We identify 15 events between April 24, 2007 and January 15, 2009 that affected the likelihood of IFRS adoption in the U.S. We use the cumulative three-day market-adjusted return centered on each event date for a sample of U.S. firms to capture investors' reactions to these events. Indicator variables based on SIC codes are assigned to identify whether the firms are in the insurance, extractive, or high-litigation-risk industries. We also expect investors in industries where IFRS is most commonly used compared to other internationally used local standards to benefit from convergence to a greater extent, since the potential reduction in information-processing costs is presumed to be larger for such industries (SEC 2008). Finally, we identify whether a firm applies LIFO and whether it operates in countries that apply IFRS to examine the potential costs and cost reductions associated with IFRS adoption.

Overall, we find a positive market reaction to the events that increase the likelihood of adoption. We also find that the positive reaction is stronger if the adoption of IFRS is expected to result in convergence benefits, and weaker for firms with high litigation risk. However, the findings do not show that investors in insurance or extractive firms are concerned about the lack of industry-specific guidance, which is inconsistent with concerns put forward by the SEC and the Financial Accounting Standards Board (FASB). In addition, we do not find that investors react more positively to IFRS adoption events if cost reductions are expected and the market reaction is not lower for LIFO firms. Our results are consistent with the view that convergence benefits matter

[1] We focus on market reactions for two reasons. First, since the switch to IFRS was proposed by the SEC and its foremost mission is to protect investors, it makes sense to examine the benefits and costs from the investors' viewpoint. This was explicitly stated by the SEC's chief accountant, James Kroeker, at the 2009 AICPA Conference. He stressed that "the fundamental focus of our evaluation of implementing a set of high-quality international standards must be on the impact to investors. I believe that implementing a single set of global accounting standards for U.S. issuers *can, and must*, be done only in a manner that is beneficial to U.S. capital markets and consistent with the SEC's mission of protecting investors" (Kroeker 2009, emphasis in the original). Second, since IFRS has not yet been adopted in the U.S., we cannot examine the direct effects of the standards on financial reporting outcomes.

[2] A "lower" market reaction indicates a less positive or more negative reaction, whereas a "higher" market reaction indicates a more positive or less negative reaction, i.e., we are referring to the algebraic direction and not the absolute magnitude of the impact.

to investors, and that the lower implementation guidance under IFRS appears to be an issue for investors in high-litigation-risk firms.

Although the study is subject to several caveats, such as the correct identification of events and the assumption that investors respond rapidly to events, the findings are relevant to the current debate on whether the SEC should move forward with the transition to IFRS, especially given the scarcity of empirical evidence to guide this decision.[3] The paper also contributes to the recent literature on the economic consequences of IFRS adoption (e.g., Barth et al. 2008; Daske et al. 2008; Armstrong et al. 2010) and provides evidence on the importance of convergence to investors.

Next, Section II offers an overview of the events that affect the likelihood of IFRS adoption in the U.S. The theoretical background is presented in Section III, Section IV discusses the sample and variables, and the main results are presented in Section V. Section VI discusses the sensitivity analyses, and Section VII provides concluding remarks.

II. EVENT HISTORY

In Spring 2007, the SEC announced for the first time that it was contemplating allowing U.S. companies to use IFRS instead of U.S. GAAP. The SEC was motivated by a longstanding desire to move to a single set of high-quality global accounting standards and by the widespread adoption of IFRS in almost 120 countries to date (IASB 2011). The underlying argument was that investors would benefit from such a move; for example, it would decrease the costs of comparing financial reports on a global basis. However, previous studies suggest that investors might not benefit significantly from this move (e.g., Hail et al. 2010), and there is little empirical evidence to substantiate the SEC's claims. This paper provides such evidence by examining U.S. investors' reactions to events that affect the likelihood of IFRS adoption in the U.S. This methodology has also been used in previous studies to assess the perceived net benefits or costs of new regulations for investors, including Christensen et al. (2007), Zhang (2007), and Armstrong et al. (2010).

To identify the events, we searched the websites of the SEC, FASB, and International Accounting Standards Board (IASB) for relevant press releases, announcements, and meetings. For related news, we searched Factiva and LexisNexis Academic Universe with the words "IFRS" and "U.S." Table 1 shows the resulting list of 15 events occurring between April 24, 2007, when the SEC first announced plans to potentially allow the use of IFRS reporting for U.S. firms, and January 15, 2009, when SEC Chairperson Mary Schapiro publicly expressed her doubts about the IFRS plans.[4] We classify 13 events as increasing the likelihood of IFRS adoption, one event as decreasing it, and we have no directional prediction for one event.

---

[3] We are aware of only one other study that examines the impact of IFRS in the U.S. Lin and Tanyi (2010) investigate market reactions to events relating to the general acceptance and use of IFRS. However, they focus on whether investors react to events that increase the use or acceptance of IFRS (e.g., their sample also includes events that capture convergence efforts between IASB and FASB) and they investigate only comparability. In contrast, this study focuses on the impact of IFRS *adoption*, since this is the key topic of debate in the U.S., and investigates several potential consequences for investors.

[4] To the best of our knowledge, these are all of the relevant events within our sample period. We concentrate on events that are publicly observable or known, which makes it easier to attribute stock returns to these news events. Moreover, our focus is on investors' perception of IFRS *adoption* specifically and thus on actions or news that relate directly to this. This is in line with the approach by Armstrong et al. (2010), who also focus specifically on events that affect the likelihood of adoption. We do not include earlier events such as convergence efforts between IASB and FASB, since we are unsure how to interpret these in our context and they were never explicitly mentioned as related to adoption of IFRS by U.S. firms. By contrast, we did include convergence events occurring after April 24, 2007 that directly relate to the adoption of IFRS. For instance, the elimination of the reconciliation requirement is included, because this was explicitly stated to be a step toward the adoption of IFRS by an SEC spokesperson.

TABLE 1

Summary of Key Events

Event	Description of Event	Increasing/ Decreasing Likelihood of Adoption	Predicted Market Reaction if: Benefits > Costs (Costs > Benefits)
(1) April 24, 2007	SEC announces plan to allow IFRS for U.S. issuers.	Increasing	+ (−)
(2) August 7, 2007	SEC Concept Release on allowing U.S. issuers to prepare financial statements in accordance with IFRS.	Increasing	+ (−)
(3) October 24, 2007	Senate hearing about international accounting standards.	Increasing	+ (−)
(4) November 7, 2007	FAF/FASB positive response to Concept Release.	Increasing	+
	Proposals for improving IASB governance.		(−)
(5) November 15, 2007	SEC approves elimination of Form 20-F reconciliation requirements for foreign issuers using IFRS as issued by IASB.	Increasing	+ (−)
(6) December 13, 2007	SEC roundtable: Should U.S. switch to IFRS?	Increasing	+ (−)
(7) December 17, 2007	SEC roundtable: What are practical implications of switching to IFRS?	Increasing	+ (−)
(8) April 18, 2008	SEC Chairman Cox states that SEC is working on a roadmap for adoption of IFRS.	Increasing	+ (−)
(9) June 16, 2008	FAF/FASB conference on IFRS: participants voice need for firm date for IFRS adoption.	Increasing	+ (−)
(10) July 21, 2008	IASB officially publishes discussion documents on IASCF Monitoring Group.	Increasing	+ (−)
(11) August 4, 2008	SEC roundtable on performance of IFRS during subprime crisis and progress of IFRS.	Increasing	+ (−)
(12) August, 27, 2008	Outlines of roadmap discussed during open meeting on IFRS held by SEC.	Increasing	+ (−)
(13) October 13, 2008	IASB adapts IAS39.	Unsigned	?
(14) November 14, 2008	Roadmap for potential use of financial statements prepared in accordance with IFRS by U.S. issuers.	Increasing	+ (−)
(15) January 15, 2009	SEC Chairwoman Mary Schapiro expresses doubts about IFRS plans.	Decreasing	− (+)

Table 1 presents a summary of all events included in the sample and their expected impact on the likelihood of IFRS adoption.

The first event occurred on April 24, 2007, when the SEC first announced that it was considering whether U.S. issuers should switch to IFRS. The SEC had long been supportive of the use of a single set of high-quality global accounting standards and now expressed the intention to move in this direction (SEC 2007). The SEC announced a planned "Concept Release relating to issues surrounding the possibility [of] providing U.S. issuers the alternative to use IFRS." At the time, the SEC was eliminating Form 20-F reconciliations for foreign firms that prepared financial statements under IFRS as promulgated by the IASB. The SEC now decided that switching from U.S. GAAP to IFRS would be the next critical step. The second event occurred on August 7, 2007 when the SEC published the Concept Release, which discussed reporting practices within and outside the U.S., potential IFRS benefits for the U.S. capital market, and implementation issues, including the training of accountants in IFRS and whether to adopt a transition period. We classify these two events as increasing the likelihood of IFRS adoption.

The third event occurred on October 24, 2007 when the U.S. Senate Subcommittee on Securities, Insurance, and Investment held an open meeting about international accounting standards.[5] Its goal was to discuss the Concept Release and the proposed elimination of Form 20-F reconciliations. Among those who testified were IASB Chairman David Tweedie, FASB Chairman Robert Herz, and representatives of the SEC, the American Institute of Certified Public Accountants (AICPA), and the Emerging Issues Task Force (EITF). The general view was that a switch to high-quality global standards would be beneficial after IFRS had been improved in areas where it lacked standards, and the differences between IFRS and U.S. GAAP were reduced or eliminated.

In response to the Concept Release, the SEC received more than 85 comment letters that were published on its website. Among the respondents were the Big 4 accounting firms that strongly supported the use of IFRS. The fourth set of events occurred on November 7, 2007 when FASB Chairman Robert Herz expressed the joint standpoint of the FASB and the Financial Accounting Foundation (FAF) in a letter. In line with his earlier statement at the Senate hearing, Herz expressed the FASB's support for the move. He also gave specific suggestions on how to achieve improvements in IFRS and the IASB's governance and funding.[6] On the same day, the IASB trustees agreed that governance improvements were necessary. The IASB published the discussion document on the proposed changes for public comment on July 21, 2008, which marks the tenth event in our sample period.[7] This document proposed establishing a monitoring group consisting of securities regulators to oversee the actions of the IASB and approve any new appointments to the board of the International Accounting Standards Committee Foundation (IASCF) trustees. Since both the Senate hearing and the FASB/FAF response by Herz were supportive of the adoption, and the IASB started to improve its governance system, we classify all three events as increasing the likelihood of IFRS adoption.

The fifth event occurred on November 15, 2007 when the SEC finally approved elimination of the Form 20-F reconciliation. This was a key step in moving toward IFRS because the elimination

---

[5] A webcast of the meeting and the testimonies are available at: http://banking.senate.gov/public/index.cfm?FuseAction=Hearings.Hearing&Hearing_ID=a96cc028-3b6d-4996-b849-768e83af35fc

[6] The concern was the IASB's lack of accountability to a single securities regulator, similar to the FASB's accountability to the SEC. This made it possible for countries to create adapted versions of IFRS, which went against the aim of the IASB to have a single set of global standards. In addition, the IASB was funded largely by the Big 4 accounting firms and voluntary donations from around 200 companies, in contrast to the FASB being funded by public companies through SEC registration fees. This led to concerns that major IASB contributors might unduly influence the standard-setting process. Since the IASB's funding and accountability was a major issue that needed to be addressed prior to the U.S. adoption of IFRS (SEC 2008), we view the announcement of governance improvements as an event that increased the likelihood of adoption.

[7] We number our events in chronological order, not in order of when they are discussed in this section.

proposal had led the SEC to consider IFRS for domestic companies (SEC 2007, 12). The SEC's chief accountant, Conrad Hewitt, called this decision "a small but significant step in moving the U.S. to IFRS" (Hewitt 2008). We therefore classify this event as increasing the likelihood of IFRS adoption.

Events six and seven occurred on December 13 and 17, 2007, when the SEC held roundtables on IFRS to discuss the role of IFRS in the U.S. capital markets and the impact on U.S. issuers' reporting. The participants were representatives from the investment community and U.S. stock exchanges, accounting firms, underwriters, academics, and U.S. issuers. The first roundtable focused on whether the U.S. should switch to IFRS for its domestic issuers, and the second on how to structure the switch. The conclusion that emerged from both roundtables was that there would certainly be benefits in the long run, in terms of higher comparability and competition for capital, from moving to a single set of global accounting standards, which would likely be IFRS. The panelists also agreed that the U.S. should not transition to IFRS without a structured plan. This would give U.S. companies, auditors, investors, and regulators time to prepare, and also allow the IASB trustees to improve IFRS in certain areas and to improve their own organization in terms of independence, accountability, and governance. However, there was no agreement on whether to mandate or allow IFRS, or when the transition should occur. There were also concerns about jurisdictional adaptations of IFRS that were related to the IASB governance problems. Despite these disagreements, the roundtables revealed clear support for the switch. Therefore, we classify these events as increasing the likelihood of IFRS adoption.

The eighth event occurred on April 18, 2008, when SEC Chairman Christopher Cox stated in an address to the U.S. Chamber of Commerce that an official "roadmap" for the adoption of IFRS would be released later that year, with more details on how the transition would be structured. The joint FAF/FASB conference on June 16, 2008 marks the ninth event. Participants including investors, auditors, educators, and issuers voiced a need for a definite date and more information about the potential adoption of IFRS, noting that without a definite date, key players would not start preparing for the switch. Despite the ongoing financial crisis, during the SEC's roundtable on August 4, 2008 (event 11), participants discussing the performance of IFRS during the crisis, and its progress in general remained positive about an imminent transition. The twelfth event occurred on August 27, 2008 when the SEC presented outlines of the roadmap at an open meeting, and all the SEC commissioners voted in favor of opening a public discussion of the roadmap. These four events again show the SEC's commitment to IFRS and the stakeholders' desire for more certainty about the timing and planning, and we classify them as increasing the likelihood of IFRS adoption.

However, the financial institutions required to use IFRS reporting were hit particularly hard by the subprime crisis. Since banks had to state certain assets at fair value and current market prices had declined significantly during the crisis, they would have been required to impair many of their assets. Largely due to political pressure from the European Commission and finance ministers, and two days after the FASB issued its Staff Position 157-3 on the same accounting issue, on October 13, 2008 the IASB hastily adapted its fair-value accounting standard IAS 39 to give financial institutions more leeway in classifying financial assets out of fair value by designating these as no longer held for sale, enabling firms to avoid these impairment losses. Although the IASB explained the adaptation as an attempt to make IFRS and U.S. GAAP more similar, its actions were deemed damaging to its credibility as an independent standard setter, also because IASB adapted IAS 39 without the usual due process and transparent procedures (Bothwell 2009). Although this event could be classified as decreasing the likelihood of IFRS, it also resulted in convergence between U.S. GAAP and IFRS in this area. As continued convergence between the two sets of standards was also an important prerequisite for IFRS adoption (SEC 2007), it might be viewed as increasing the likelihood of adoption. Since it is difficult to predict investors' reactions, we classify this 13th event as neither increasing nor decreasing the likelihood of IFRS adoption.

Despite the controversy surrounding IAS 39, the SEC published the official roadmap on November 14, 2008, and we classify this 14th event as increasing the likelihood of IFRS adoption. The roadmap sets out seven milestones, which if attained by 2011 would likely lead to the phased mandatory adoption of IFRS by 2014, although the roadmap suggests that some companies will be eligible for early voluntary adoption.[8] Although the roadmap suggests 2014 as the adoption year, the final decision was to be made in 2011.[9]

In January 2009, SEC Chairman Cox, who was largely responsible for developing the IFRS plans, was succeeded by Mary Schapiro, who made several critical remarks about these plans. On January 15, Schapiro expressed her doubts about the IFRS plans at her confirmation hearing before the Senate Banking Committee, stating that she would not be bound by the roadmap and would take time to carefully review the plans before proceeding. She further expressed concerns regarding the IASB's lack of political independence and the quality of IFRS compared to U.S. GAAP. Since these remarks signaled the SEC's intention to delay, or even halt, the adoption process (Johnson 2009), we classify this final event as decreasing the likelihood of adoption.

III. THEORETICAL BACKGROUND

Convergence and IFRS Adoption

Whether convergence in accounting standards benefits investors is a much-debated issue. Convergence means increasing the compatibility of accounting standards while maintaining a high level of quality (Pacter 2005; Zeff 2007). For U.S. GAAP and IFRS in particular, convergence efforts have ranged from the joint efforts of the FASB and IASB to make existing standards more similar to the potential adoption of IFRS for use by U.S. companies. Regulators and standard setters often emphasize that convergence benefits investors through lower information-processing costs, since it reduces the need for investors to learn and understand different sets of accounting standards (Chi 2009). Convergence could increase the quality and comparability of financial reporting (SEC 2008; Hail et al. 2010) and enhance the consistency of financial information by enabling auditors and their clients to develop consistent global practices to deal with accounting issues (Tweedie 2006).[10] However, the extent to which these benefits will be realized is unclear. For instance, Barth

---

[8] The milestones in the roadmap relate to issues such as improvements in IFRS, in the accountability and funding of the IASCF, and in the ability to use XBRL for IFRS reporting, and training in IFRS, that must be addressed before the mandatory adoption of IFRS. They also relate to the transition plan for mandatory IFRS, including successful early use by eligible firms, the anticipated timing of future rule making by the SEC, and the implementation of mandatory IFRS. The roadmap states that if a firm is among the largest 20 in its industry based on market capitalization, and the industry's most commonly used accounting method worldwide is IFRS, the firm may be allowed to voluntarily adopt IFRS.

[9] On February 24, 2010, the SEC announced that it had changed the proposed adoption date to 2015 to allow companies more time to prepare (see http://www.sec.gov/news/press/2010/2010-27.htm). To date, the SEC has not yet made a final decision on whether the U.S. will adopt IFRS; an announcement was expected mid-2012 (see http://www.journalofaccountancy.com/Web/20125186.htm).

[10] Similar to Armstrong et al. (2010) and Hail et al. (2010), we view increased reporting quality and convergence benefits as two different but related effects of IFRS adoption. Because there is no standard definition of reporting quality, we view it as the extent to which financial reporting reflects a firm's underlying economic performance. Research has associated quality with earnings attributes such as the degree of earnings management, timely loss recognition, and value relevance (e.g., Francis et al. 2004; Barth et al. 2008) or the quantity of disclosure (e.g., Botosan 1997; Leuz and Verrecchia 2000). As explained above, convergence benefits are broader than reporting quality and can include reduced information-processing costs due to greater ease in comparing firms' financial performance globally. One potential benefit of convergence is comparability, which is the extent to which the information presented allows investors "to identify the similarities in and differences between two sets of economic phenomena" (FASB 1980). Even if reporting quality is held constant, comparability can increase the usefulness of reporting to investors by making it less costly to compare firms (Hail et al. 2010), which has been a key motivation for allowing or requiring the use of IFRS (see, e.g., FASB 2008; FAF 2009; FCAG 2009).

et al. (1999) show that conceptually, the effect of harmonization or convergence is ambiguous. Depending on its impact on the precision of GAAP and investors' costs and benefits of acquiring expertise in understanding different GAAPs, harmonization may not always lead to more precise information and capital market benefits. In addition, there are different views on whether uniformity in accounting standards is desirable. On the one hand, the SEC has long supported global convergence in accounting standards (SEC 2007) and Barth (2008) states that the use of a common reporting language in business, or a single set of accounting standards, is an important step in making financial reporting more comparable. However, opponents argue that convergence may not leave room for "considering differences in circumstances among companies or countries" (Zeff 2007) and could even result in less informative reporting if a "one-size-fits-all" approach obscures underlying performance or characteristics of firms and thus could result in a loss of information (Chi 2009). Moreover, Kothari et al. (2010) predict that forcing FASB and IASB to compete instead of converge would lead to GAAP rules that facilitate efficient capital allocation. Finally, prior research also highlights the importance of reporting incentives together with accounting standards (e.g., Hung 2000; Ball et al. 2000, 2003; Ball and Shivakumar 2005; Burgstahler et al. 2006), meaning that convergence alone may not necessarily result in more informative reporting and capital market benefits.

Empirically, findings from prior literature provide evidence that convergence does result in capital market benefits and changes in financial reporting characteristics. For instance, Chi (2009) examines whether investors' ability to process earnings information is hindered by firms' use of different GAAPs. She finds that when multiple firms announce their earnings on the same day, the price and trading-volume reaction is greater and the post-earnings-announcement drift is smaller if these firms use fewer different domestic GAAPs. This suggests that investors are able to process information more efficiently when the analysis is not complicated by the presence of multiple standards, which supports convergence as being beneficial to capital markets. Other studies on the effects of IFRS adoption in particular have shown that IFRS results in greater reporting quality and requires greater disclosure than most local GAAPs (Ashbaugh and Pincus 2001; Barth et al. 2008), and can result in greater reporting comparability (Yip and Young 2011). Theoretical research shows that this can reduce information asymmetry problems and estimation risk, which in turn has benefits for liquidity and the cost of equity (Diamond and Verrecchia 1991; Lambert et al. 2007).

Armstrong et al. (2010) find empirical support for this prediction in a European setting, where share prices react positively to events that increase the likelihood of IFRS adoption, in particular for firms that are expected to benefit from IFRS in terms of higher information quality and convergence. Beuselinck et al. (2011) find that disclosure under IFRS revealed new firm-specific information in the year of mandatory adoption in the EU, which subsequently reduced the surprise of future disclosures. There is also evidence that mandatory IFRS adopters experience improvements in liquidity, cost of capital, and equity valuation (Daske et al. 2008). Drake et al. (2010) find that these increases in market liquidity are greater for firms with high-quality pre-adoption information environments. Since these firms are unlikely to benefit from increased reporting quality, Drake et al. (2010) attribute these positive market effects to increased comparability. Li (2010) also shows that mandatory IFRS adopters experience a decrease in cost of equity and that this can be attributed in part to increased comparability as well as to greater disclosure under IFRS. Wu and Zhang (2010) find that in the EU the use of relative performance evaluation with international industry peers increases after IFRS adoption, while DeFond et al. (2011) report increased U.S. mutual fund ownership in firms that credibly adopt IFRS, which they interpret as consistent with increased comparability under IFRS. Furthermore, in line with prior research, the cited studies show that the effect of IFRS is highly dependent on reporting incentives shaped by regulatory enforcement and other institutional factors.

Potential Effects of IFRS Adoption in the U.S.

Although prior studies find positive capital market effects associated with convergence and IFRS adoption in particular, these findings do not necessarily apply to the U.S. context. First, there are opposing views on whether IFRS is, overall, of higher quality than U.S. GAAP (Hail et al. 2010). IFRS proponents argue that it is less complex than U.S. GAAP, and that the nature of current U.S. standards induces managers to follow rules rather than to consider whether financial reporting reflects the underlying economics of a firm. On the other hand, critics of IFRS claim that its principles-based nature can be abused by managers, since more discretion and less guidance leave more room for earnings management. Also, IFRS and U.S. GAAP have become increasingly similar over time, as the FASB and IASB have worked together intensively to increase and maintain the compatibility of standards (Hail et al. 2010). Examples include IASB's new standards on borrowing costs (IAS23R) and segment reporting (IFRS8) that mirror U.S. GAAP.

If investors believe that these convergence efforts have sufficiently reduced the differences, then adopting IFRS would not result in significant convergence benefits and would bias against finding a more positive market reaction. However, anecdotal evidence suggests that the application of U.S. GAAP versus IFRS still results in different reporting outcomes. For example, Ahold, a Dutch food retailer that operates internationally, showed a net profit of €120 million for 2005 under IFRS, but reported a net loss of €20 million for the same year under U.S. GAAP in its reconciliation footnote. This illustrates that despite ongoing convergence, the use of different accounting standards has a material impact on financial reporting. Second, since reporting quality in the U.S. is among the highest in the world (Leuz et al. 2003), and factors such as incentives play an important role in determining this quality, it is unclear whether the adoption of IFRS will have a significant impact on the quality of reporting in the U.S.

We acknowledge that it is difficult to predict the overall effect of IFRS adoption. However, we expect cross-sectional variation in the extent to which it is beneficial or costly. We therefore focus on three settings where the effects of adopting IFRS are expected to be most pronounced. First, we examine whether IFRS adoption is perceived by investors as more costly in industries where it will most likely decrease the quality of standards. Although U.S. GAAP and IFRS have become increasingly similar (Hail et al. 2010), the SEC and FASB have expressed concerns about the lack of IFRS implementation guidelines for certain industries, notably the extractive and insurance industries. Their concern is that investors might lose information that is currently available under U.S. GAAP. To the extent that the lack of industry-specific guidance is indeed a concern, then investors in these firms might be opposed to IFRS adoption and react negatively to events that increase the likelihood of adoption. Also, the lack of specific rules could be problematic for industries with high litigation risk. Managers will have to rely more on their own judgment when interpreting IFRS, which could result in more legal challenges to their decisions. To avoid this, firms might make overly conservative accounting choices (Hail et al. 2010) that reduce the informativeness of financial reporting. If the lack of implementation guidance is indeed viewed as a valid concern by investors, we would expect to observe a less positive market reaction for firms in extractive, insurance, and high-litigation-risk industries.

Second, we expect convergence benefits to be more pronounced in industries where many non-U.S. peer firms have already adopted IFRS. Widespread adoption of IFRS in a particular industry may be an indication that the benefits (such as reduced information-processing costs) of adopting these standards are greater, resulting in a larger proportion of non-U.S. firms adopting IFRS. Analogously, these benefits may also apply to U.S. firms in such industries, thus resulting in a more positive market reaction to IFRS adoption events for these firms. In line with this argument, the fact that many global competitors use IFRS would indicate more consistent global practices to deal with accounting issues, and possibly greater familiarity among the international investment community

with IFRS reporting in that industry. Supporting this view, the SEC considered allowing certain U.S. firms for which IFRS would be most beneficial to adopt IFRS early, and proposed that the use of IFRS by a majority of significant competitors should be the key requirement for deciding which firms would be eligible for this option (SEC 2008). For these reasons, we expect that U.S. firms in such industries would benefit from IFRS adoption to a greater extent than firms in industries where IFRS is not widely adopted by non-U.S. peers.

Third, we examine the potential costs and cost reductions of IFRS adoption. Experience with the adoption of new accounting regulations has shown that there are substantial implementation costs. For example, the implementation of SOX Section 404 costs an estimated \$3 to \$8 million per firm (FEI 2004). The Institute of Chartered Accountants in England and Wales (ICAEW) issued a report discussing the compliance costs of IFRS adoption in Europe. They estimated these costs to be between 0.05 percent (for larger companies) and 0.31 percent (for smaller companies) of revenue. The SEC's estimate of implementation costs for the largest U.S. firms is around 0.125 percent of revenue, or around \$32 million per firm for the first three years of filings on Form 10-K (SEC 2008).

IFRS could also result in a recurrent loss of tax benefits for firms that use LIFO. Since U.S. tax regulations require the use of LIFO for tax-reporting purposes, and IFRS does not permit the use of this method, firms applying LIFO valuation would be forced to forgo tax benefits.[11] Although Hail et al. (2010) suggest several approaches to this issue, such as dropping the book-tax conformity requirement or providing a tax credit to LIFO firms, investors in LIFO firms might react negatively to IFRS if it results in substantially higher taxes and lower cash inflows.

By contrast, U.S.-based multinationals might benefit from recurrent cost reductions. Corporations with subsidiaries in countries with mandatory IFRS reporting may be able to reduce their costs because they no longer have to report under both U.S. GAAP and IFRS (SEC 2008; Hail et al. 2010). Investors in these multinationals might therefore react positively to IFRS adoption.

Predictions

Based on the above discussion, we make the following cross-sectional predictions regarding market reactions. First, investors are likely to be concerned that IFRS will adversely affect reporting quality because of the lack of implementation guidance for extractive and insurance firms and firms with high litigation risk. We therefore expect a less positive market reaction for such firms. Second, we expect investors to react more positively if they expect IFRS to result in convergence benefits, which is more likely in industries where IFRS is already widely adopted by non-U.S. peer firms. Finally, we examine whether investors consider the cost impact of IFRS in their valuations. IFRS is expected to decrease costs for firms that must comply with both U.S. GAAP and IFRS, so we expect such firms to experience a most positive market reaction. By contrast, we expect a lower reaction for firms that use LIFO because of the higher costs resulting from a loss of tax benefits.

IV. SAMPLE SELECTION, VARIABLES, AND DESCRIPTIVE STATISTICS

Sample Selection

Because we are mainly interested in the cross-sectional differences in investor responses, the sample includes all domestic U.S. firms that have the necessary price and financial statement data

[11] Although this may seem a minor concern, more than 120 U.S. companies have joined the LIFO Coalition, which aims to preserve the use of LIFO, and have lobbied against the adoption of IFRS (see http://www.sec.gov/comments/s7-27-08/s72708-45.pdf). It is ultimately an empirical question whether investors consider the potential loss of tax benefits to be economically significant.

for the period 2007 to 2009 encompassing the 15 events. Firms are not required to have data for all 15 events, i.e., they are included if they have return data for at least one event and data for the corresponding variables in the cross-sectional analyses. We start with 63,597 domestic U.S. firm-event observations with return data for one of the 15 events. We do not exclude any industries, but do exclude observations for firms with a negative book value of equity, resulting in 61,610 observations. We further exclude firms for which we lack data to calculate the test or control variables, resulting in the final sample of 59,285 firm-event observations for 4,820 firms. Details on sample selection are provided in Appendix A.

Variable Definitions

The variables used in the analyses are discussed in this section. Further details on the variable measurement and data sources are provided in Appendix B.

Dependent Variable

The dependent variable is the market reaction to the events identified in Section II. This is measured by the three-day cumulative return centered on the event date, retrieved from CRSP, and is adjusted for other news using a market index. Because IFRS adoption would affect all publicly listed U.S. firms, it is inappropriate to adjust returns with a U.S.-based index, because the index itself would also reflect the market reaction to IFRS adoption events. In the spirit of Armstrong et al. (2010), we use the three-day return to the DJ STOXX 1800 index, excluding American firms as the market adjustment for the main analyses.[12] We obtain the return data for the STOXX adjustment from Datastream.

Test Variables

First, as discussed in Section III, in some instances reporting quality may decrease due to less implementation guidance. We define indicator variables that capture whether a firm operates in the insurance (*INSUR*) or extractive (*EXTR*) industries, or in industries with a high litigation risk (*HI-LIT*). *INSUR* is equal to 1 if a firm's two-digit SIC code (*SIC2*) is 63 or 64, and *EXTR* is equal to 1 if *SIC2* is 13 or 29.[13] Following Kasznik and Lev (1995), Matsumoto (2002), and Field et al. (2005), *HI-LIT* is equal to 1 if a firm's SIC is 2833–2836, 3570–3577, 3600–3674, 5200–5961,

---

[12] Prior studies that have also used non-U.S. firms' returns to adjust for other contemporaneous economic news include Zhang (2007) and Armstrong et al. (2010). A potential drawback of using this DJ STOXX 1800 index excluding American firms is that it includes European firms, most of which will have adopted IFRS. To the extent that these firms also respond positively to IFRS adoption events in the U.S. since it could affect the comparability of their financial reporting, this adjustment could remove some of the effect of the IFRS adoption news we seek to document. An alternative would be to use the DJ STOXX Asia-Pacific index returns, which consist of the 600 largest firms in the Asia-Pacific. One drawback of this index is that Asian and U.S. firms may be dissimilar, so adjusting with this index may not adequately remove market reactions to news unrelated to IFRS adoption. We base our choice of adjustment on the correlations between the index returns and our sample returns. For 2007–2008, the correlation between three-day U.S. returns and the STOXX 600 Asia-Pacific index is 0.64, and the correlation between the U.S. returns and the STOXX Global 1800 ex America index is 0.79. On the event dates, the correlations are 0.66 for the former index and 0.88 for the latter. For completeness' sake, Table 3 shows results for both indices, and Table 6 reports regression results for the STOXX Asia-Pacific index. Nevertheless, we acknowledge the challenge of selecting an appropriate market adjustment, and that using non-U.S. firms' returns as an adjustment does not remove the impact of news events that are unrelated to IFRS and specific to the U.S. (Leuz 2007).

[13] The industry classification of insurance companies is based on Fama and French (1997) and that of extractive industries is based on Hand (2003).

7371–7379, or 8731–8734. If investors expect IFRS adoption to lead to lower reporting quality, we expect the coefficients on these three variables to be negative.

Second, we test whether investors respond more positively to IFRS adoption events if IFRS is expected to result in convergence benefits. As explained earlier, we expect these benefits to be most pronounced in industries where a majority of firms apply IFRS. To assess this, we look at the accounting standards that are applied by peer firms in the same industry on a worldwide basis. We use the Worldscope database to determine the accounting standards used by peers.[14] Global industry peers are selected from Worldscope by ranking firms on market capitalization in two-digit SIC groups.[15] For the 20 largest firms in each industry group, we determine the statistical mode of the accounting standards and classify an industry as "IFRS-predominant" if IFRS is the most commonly used standard among these 20 firms where we set $D(IFRS) = 1$. We expect a positive coefficient for *D(IFRS)* if investors expect IFRS adoption to lead to net convergence benefits for these firms. In Section VI we conduct several additional analyses to gain more insight into the nature of these convergence benefits.

Third, we examine whether the market reactions vary with the cost impact of IFRS. We expect U.S. firms operating in countries with mandatory IFRS reporting to experience reduced costs by switching to IFRS. We first use the Compustat segment files to determine the geographic origin of a firm's sales. We then determine which countries have mandated the use of IFRS, using the IASPlus website and the classification by Sletten and Ramanna (2009). The variable *IFRS SALES%* represents the proportion of sales in countries that mandate IFRS relative to the firm's total sales. We expect firms with a higher proportion of IFRS sales to benefit more from IFRS adoption. Finally, we determine whether a company uses LIFO via Compustat information about inventory-valuation methods. LIFO is an indicator variable equal to 1 if a firm applies this method to value its inventory.

Control Variables

Following Christensen et al. (2007) and Armstrong et al. (2010), we include the following additional control variables: firm size, turnover, leverage, industry concentration, and auditor size. These variables are proxies for the firm's information environment and information asymmetry with investors. On the one hand, smaller and less liquid firms with low turnover, firms with higher leverage, and those in more concentrated industries are expected to have poorer information environments. If investors expect IFRS adoption to improve reporting quality and reduce information asymmetry, we would expect such firms to benefit more from the adoption of IFRS. On the other hand, larger and more liquid firms attract more institutional owners who prefer conformity in accounting choices (Bradshaw et al. 2004). If investors expect IFRS adoption to result in such convergence benefits, we would expect a more positive reaction for companies where the demand for conformity is higher. Firms with a Big 4 auditor are also expected to benefit more from IFRS adoption, since such auditors are better equipped to support the transition (Armstrong et al. 2010).

---

[14] Daske et al. (2008) show that some firms have incorrect accounting-standard classifications in Worldscope. We acknowledge that there are flaws in the commercially available databases, but we do not think that this is a severe problem for our study, since our data are derived from a more recent period and we use only the largest firms to determine the most common standards. Misclassification is likely to be a more serious issue in the years preceding mandatory IFRS adoption and for smaller firms.

[15] We also conduct analyses with different industry groupings based on the three-digit SIC as well as the Industry Classification Benchmark (ICB) subsector codes. The ICB system was developed by Dow Jones Indexes and FTSE and is also used in Lang et al. (2010). ICB classifies firms into industry sectors based on their sources of revenue. This system has been adopted by many stock exchanges globally and aims to offer a comprehensive tool for global sector analysis, with a focus on relevance to investors. The findings are robust to alternative industry definitions and are discussed in more detail later in the paper.

Descriptive Statistics and Correlations

Table 2 presents descriptive statistics and correlations. Panel A shows that the proportion of firms in extractive (insurance) industries is 4.1 percent (3.3 percent), and 25.12 percent of firms are in high-litigation industries. Furthermore, 40.01 percent of firms operate in an industry where IFRS is commonly used internationally.[16] The descriptive statistics for *D(IFRS SALES)*, which equals 1 if a firm has non-zero sales in a country that mandates IFRS, indicate that 19.80 percent of firms generate sales in a country that mandates IFRS reporting. For such firms, the average proportion of IFRS sales is 22.15 percent of total sales. Additionally, 8.08 percent of sample firms apply LIFO valuation.

Panel B reveals that the correlation between the market reaction to events ($CR_{STOXX\ ex\ Am.}$) and *EXTR* is positive and significant at the $p < 0.05$ level, which is inconsistent with the view that investors in extractive industries are concerned about the lack of industry-specific guidance under IFRS. The same holds for the insurance industries (*INSUR*). By contrast, *HI-LIT* is negatively correlated with the market reaction, consistent with the notion that less guidance under IFRS is perceived as costly for firms with a high litigation risk. The correlation between $CR_{STOXX\ ex\ Am.}$ and *D(IFRS)* is also positive, consistent with convergence benefits being greater in IFRS-predominant industries. The correlation between the event returns and the variables *D(IFRS SALES)* and *IFRS SALES%* is insignificant, which does not support the prediction that investors react more positively to IFRS adoption because of reduced reporting costs. The correlation between *LIFO* and $CR_{STOXX\ ex\ Am.}$ is positive and significant, suggesting that investors do not perceive IFRS adoption to be more costly due to the potential loss of tax benefits. Overall, the correlations in Panel B are preliminary evidence that IFRS is perceived as costly for firms with high litigation risk and as beneficial in cases where investors expect IFRS to lead to convergence benefits.

V. RESULTS

Overall Market Reaction

We first examine the overall market reaction to the 15 events to assess whether U.S. investors on average perceive IFRS adoption to be net beneficial or costly. Table 3 shows an average positive abnormal return across all events (adjusting for STOXX 1800 ex America) of 0.86 percent, which is marginally statistically significant (t-statistic = 1.94; two-tailed p-value = 0.0724). The significance is determined using the empirical distribution of the value-weighted returns for the 15 events, assuming that the mean returns per event are uncorrelated across events.[17] We use the mean (–0.0006) of non-event returns adjusted for STOXX 1800 ex America measured over non-overlapping three-day windows as the benchmark, rather than assuming $H_0 = 0$ as explained in Appendix B. This allows for unequal variances between event and non-event return distributions and does not assume that the market adjustment fully adjusts for the market return (Armstrong et al. 2010). This second t-statistic is slightly higher (2.08) and marginally statistically significant with a two-tailed p-value of 0.056. We find similar and statistically stronger results with the STOXX Asia-Pacific adjustment.

---

[16] IFRS is predominantly used in the agricultural and mining industries, several manufacturing industries (such as food, fabrics, and electronic equipment), trucking and transportation, finance and insurance, and leisure services.

[17] To calculate the t-statistics in Table 3 and for the cross-sectional analyses in Tables 4 to 6, we multiply the returns for event 15, which is classified as decreasing the adoption likelihood by –1. This is done to ease the interpretation of the t-statistics, since all the other events are classified as increasing adoption likelihood or are unsigned (Armstrong et al. 2010).

TABLE 2

Descriptive Statistics

Panel A: Distribution of Variables

Variable	Mean	25%	Median	75%	Std.
$CR_{STOXX\ ex\ Am.}$	0.0052	−0.0263	0.0034	0.0342	0.0688
$CR_{STOXX\ A.P.}$	0.0084	−0.0274	0.0069	0.0418	0.0711
EXTR	0.0413	0	0	0	0.1991
INSUR	0.0326	0	0	0	0.1775
HI-LIT	0.2512	0	0	1	0.4337
D(IFRS)	0.4001	0	0	1	0.4899
D(IFRS SIMILAR)	0.1980	0	0	0	0.3985
D(IFRS SIC3)	0.1613	0	0	0	0.3678
D(IFRS ICB)	0.3748	0	0	1	0.4841
ANALYST	5.3594	1	4	8	5.6724
# FOREIGN INSTITUTIONAL OWNERS	10.7341	2	5	11	15.0068
D(IFRS SALES)	0.1961	0	0	0	0.3971
IFRS SALES%	0.2215	0.1057	0.1832	0.2900	0.1701
LIFO	0.0808	0	0	0	0.2726
SIZE	13.1055	11.7442	13.0042	14.3441	1.9274
HERFINDAHL	0.0748	0.0347	0.0509	0.0800	0.0725
LEVERAGE	0.5299	0.3218	0.5308	0.7473	0.2598
TURNOVER	0.5237	0	1	1	0.4994
BIG4	0.6884	0	1	1	0.4632

(continued on next page)

TABLE 2 (continued)

Panel B: Correlations

	Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]
[1]	$CR_{STOXX\ ex\ Am.}$		**0.02**	**0.03**	**−0.03**	**0.03**	**0.05**	**0.05**	−0.01	−0.01	**0.01**	**0.08**	0.00	**0.02**	**0.03**	**0.05**
[2]	*EXTR*	**0.03**		**−0.04**	**−0.12**	**−0.16**	**0.12**	**0.04**	**−0.04**	**−0.02**	0.01	**0.07**	**−0.05**	**−0.04**	**0.06**	**−0.02**
[3]	*INSUR*	**0.02**	**−0.04**		**−0.11**	**0.15**	**0.04**	**0.08**	**−0.08**	**−0.06**	**−0.05**	**0.10**	**−0.01**	**0.12**	**−0.05**	**0.07**
[4]	*HI-LIT*	**−0.03**	**−0.12**	**−0.11**		**−0.39**	**0.13**	0.00	**0.08**	**0.06**	**−0.05**	**−0.05**	**−0.14**	**−0.29**	**0.15**	**0.06**
[5]	*D(IFRS)*	**0.02**	**−0.16**	**0.15**	**−0.39**		**−0.14**	**−0.02**	**−0.22**	**−0.18**	**−0.07**	**0.02**	0.00	**0.44**	**−0.22**	**−0.09**
[6]	*ANALYST*	**0.07**	**0.09**	**0.04**	**0.11**	**−0.14**		**0.70**	**0.11**	**0.05**	**0.11**	**0.73**	**0.03**	**−0.01**	**0.40**	**0.42**
[7]	*# FOREIGN INSTOWN*	**0.10**	**0.05**	**0.08**	**0.00**	**−0.04**	**0.71**		**0.14**	**0.08**	**0.23**	**0.81**	**0.05**	**0.10**	**0.25**	**0.37**
[8]	*D(IFRS SALES)*	0.01	**−0.04**	**−0.08**	**0.08**	**−0.22**	**0.12**	**0.15**		**0.76**	**0.13**	**0.13**	**−0.05**	**−0.16**	**0.11**	**0.12**
[9]	*IFRS SALES%*	0.01	**−0.04**	**−0.08**	**0.08**	**−0.22**	**0.12**	**0.15**	**0.99**		**0.07**	**0.07**	**−0.05**	**−0.11**	**0.08**	**0.07**
[10]	*LIFO*	**0.02**	0.01	**−0.05**	**−0.05**	**−0.07**	**0.12**	**0.20**	**0.13**	**0.13**		**0.23**	**0.15**	**0.02**	**0.05**	**0.15**
[11]	*SIZE*	**0.10**	**0.08**	**0.10**	**−0.05**	**0.01**	**0.76**	**0.88**	**0.13**	**0.13**	**0.22**		**0.06**	**0.10**	**0.36**	**0.54**
[12]	*HERFINDAHL*	0.00	**−0.10**	**0.01**	**−0.27**	**0.11**	**0.02**	**−0.03**	**−0.07**	**−0.07**	**0.14**	0.01		**0.03**	**−0.02**	**0.02**
[13]	*LEVERAGE*	**0.02**	**−0.04**	**0.12**	**−0.29**	**0.45**	−0.01	**0.05**	**−0.16**	**−0.15**	**0.02**	**0.10**	**0.19**		**−0.15**	**−0.02**
[14]	*TURNOVER*	**0.04**	**0.06**	**−0.05**	**0.15**	**−0.22**	**0.47**	**0.47**	**0.11**	**0.11**	**0.05**	**0.41**	**−0.05**	**−0.15**		**0.34**
[15]	*BIG4*	**0.06**	**−0.02**	**0.07**	**0.06**	**−0.09**	**0.50**	**0.54**	**0.12**	**0.11**	**0.15**	**0.57**	**−0.04**	**−0.01**	**0.34**	

This table presents descriptive statistics for variables used in cross-sectional analyses. For both panels, n = 59,285, except for *IFRS SALES%* in Panel A, where n = 11,628. Panel A presents distributions; correlations are presented in Panel B (Pearson (Spearman) correlations are above (below) diagonal). Distributional statistics of *IFRS SALES%* in Panel A are for firms that have non-zero sales in IFRS countries. Correlations in bold are significant at the 5 percent level or less for two-tailed tests. Variable definitions are given in Appendix B.

TABLE 3
Overall Reaction to Events Affecting Likelihood of IFRS Adoption

Event Date	Impact on Adoption Likelihood	*CR*	S&P 500 Index	*STOXX 1800 ex America*	$CR_{STOXX\ ex\ Am.}$	*STOXX 600 Asia-Pacific*	$CR_{STOXX\ A.P.}$	VIX
April 24, 2007	Increasing	0.0065	0.0075	−0.0043	0.0108	−0.0099	0.0164	13.12
August 7, 2007	Increasing	0.0443	0.0444	0.0172	0.0271	0.0038	0.0405	21.98
October 24, 2007	Increasing	0.0039	0.0054	0.0135	−0.0096	0.0055	−0.0016	20.79
November 7, 2007	Increasing	−0.0162	−0.0179	−0.0116	−0.0046	−0.0214	0.0052	24.68
November 15, 2007	Increasing	−0.0161	−0.0151	−0.0041	−0.0121	0.0057	−0.0219	26.50
December 13, 2007	Increasing	−0.0103	−0.0065	−0.0262	0.0159	−0.0430	0.0328	22.77
December 17, 2007	Increasing	−0.0229	−0.0225	−0.0214	−0.0014	−0.0361	0.0132	23.48
April 18, 2008	Increasing	0.0136	0.0172	0.0216	−0.0080	0.0387	−0.0252	20.33
June 16, 2008	Increasing	0.0099	0.0084	0.0139	−0.0039	0.0243	−0.0144	21.10
July 21, 2008	Increasing	0.0145	0.0132	0.0175	−0.0029	0.0211	−0.0066	22.76
August 4, 2008	Increasing	0.0072	0.0142	−0.0175	0.0247	−0.0394	0.0466	22.40
August 27, 2008	Increasing	0.0289	0.0265	0.0081	0.0209	−0.0089	0.0378	19.89
October 13, 2008	Unsigned	0.1070	0.0987	0.0601	0.0470	0.0607	0.0463	60.02
November 14, 2008	Increasing	−0.0014	0.0017	−0.0252	0.0238	−0.0361	0.0347	65.10
January 15, 2009	Decreasing	−0.0321	−0.0246	−0.0300	−0.0020	−0.0063	−0.0258	48.75
Mean return across events		0.0134	0.0133	0.0048	0.0086	−0.0019	0.0153	
t-statistic regular					1.94*		2.39**	
(two-sided p-value)					(0.0724)		(0.0318)	
Non-event average					−0.0006		−0.0040	
t-statistic non-event					2.08*		3.00***	
(two-sided p-value)					(0.0563)		(0.0095)	
Mean return excluding event 13		0.0067	0.0072	0.0008	0.0059	−0.0064	0.0131	
t-statistic regular					1.57		2.02*	
(two-sided p-value)					(0.1408)		(0.0641)	
t-statistic non-event					1.73		2.64**	
(two-sided p-value)					(0.1072)		(0.0206)	

*, **, *** Indicate significance at 0.10, 0.05, and 0.01 levels, respectively.

Table 3 presents the value-weighted mean return by event. We multiply returns for event 15 by −1 to calculate the mean across events (see footnote 17). Variables are as defined in Appendix B. S&P 500 Index is three-day cumulative return for S&P 500 index. VIX is the three-day average Chicago Board Options Exchange (CBOE) volatility index. *t-statistic regular* shows significance of mean return across events with $H_0 = 0$. For *t-statistic non-event*, $H_0 =$ mean of non-overlapping STOXX-adjusted three-day non-event returns over 2007 and 2008.

For comparison purposes, we also include the three-day raw returns for the S&P 500 index for each of the 15 events alongside the three-day raw returns of our sample firms in Table 3, to allow readers to assess the representativeness of the sample. We do not find major discrepancies between the returns for our sample and the S&P 500 index in terms of direction or magnitude.

Although these returns appear to indicate that the overall market reaction to events that increase the likelihood of adoption is positive, we do not interpret this as evidence for the overall desirability of IFRS adoption. We acknowledge the need to appropriately adjust stock returns and to control for the effect of confounding events. This is especially important because the period of interest coincides with the financial crisis and the heightened volatility of financial markets. The last column of Table 3 shows the Chicago Board Options Exchange Volatility index over the three days of each event window. We find that volatility is especially high during the last three events in our sample. In particular, event 13 coincides with the stock market crash of 2008, which had a large influence on the returns during that event. Immediately preceding this event, the S&P 500 had lost 22 percent of its value over the course of six trading days of October 2 to October 6 (Steverman 2008), while the Dow Jones Industrial Average (DJIA) fell 18 percent in the week starting October 6, making it the worst week in the history of the index (Curran 2008). On Monday October 13, global stock markets temporarily recovered as governments announced plans to bail out financial institutions. These extreme conditions make it difficult to interpret the return on this particular date and it is unlikely that it reflects investors' reaction to IFRS adoption news alone, but it is unclear if and how it would influence the cross-sectional results. In our opinion, the main contribution lies in the cross-sectional analyses presented below, since these results allow for more rigorous testing of alternative explanations for our tests. We find that the results are generally robust to excluding event 13 or any of the 15 events, and discuss this issue in more detail in Section VI.

Cross-Sectional Analyses

This section presents the results from cross-sectional analyses that examine whether market reactions vary across firms according to our theoretical predictions. We estimate the following model, which includes all test variables and control variables simultaneously:

$$CR_{STOXX\ ex\ Am.\ i,\ e} = f\Big(EXTR_{i,e},\ INSUR_{i,e},\ HI\text{-}LIT_{i,e},\ D(IFRS)_{i,e},\ IFRS\ SALES\%_{i,e}, LIFO_{i,e},\ control\ variables\Big), \qquad (1)$$

where i denotes firm i and e denotes the event.

We recognize that news of IFRS adoption affects all firms in the sample simultaneously, potentially resulting in cross-sectional correlations in returns and biased standard errors (Petersen 2009). To address this concern, the reported t-statistics (in parentheses below the coefficients) are based on standard errors clustered at the event level and are adjusted for heteroscedasticity.

Considering the results in the first column of Table 4, we find no support for the idea that investors in extractive and insurance industries respond more negatively to IFRS adoption events. The coefficients for *EXTR* and *INSUR* are positive but insignificant. This result could reflect investors' confidence in the efforts of the IASB to develop specific standards for these two industries. In particular, this would include the second phase of the comprehensive insurance contracts project to replace the current IFRS 4, and the efforts aimed at developing a new standard considering all unique issues of the extractive industry, to replace IFRS 6.

Second, we find a significantly negative coefficient for the *HI-LIT* variable. This indicates that investors in firms with high litigation risk react more negatively to events that increase the likelihood of IFRS adoption. This is consistent with concerns that investors may have about the lack of specific guidance under IFRS resulting in higher litigation risk.

TABLE 4

Main Cross-Sectional Analyses

Panel A: Cross-Sectional Analyses

$CR_{STOXX\ ex\ Am.} = f(EXTR, INSUR, HI\text{-}LIT, CONVERGENCE, IFRS\ SALES\%, LIFO, Control\ Variables)$.[a]

	Prediction	(1)	(2)	(3)	(4)
Intercept	?	−0.0261	−0.0209	−0.0225	−0.0237
		(−2.58)**	(−2.50)**	(−2.51)**	(−2.50)**
		[−1.99]*	[−1.75]	[−1.87]*	[−1.84]*
EXTR	−	0.0112	0.0119	0.0096	0.0117
		(1.52)	(1.56)	(1.26)	(1.56)
		[1.73]	[1.79]	[1.45]	[1.71]
INSUR	−	0.0049	0.0037	0.0064	0.0068
		(1.00)	(0.81)	(1.23)	(1.25)
		[0.84]	[0.62]	[1.08]	[1.15]
HI-LIT	−	−0.0023	−0.0021	−0.0028	−0.0020
		(−2.14)**	(−1.91)**	(−2.36)**	(−1.58)*
		[−1.70]*	[−1.40]*	[−1.99]*	[−1.40]*
CONVERGENCE[a]	+	0.0030	0.0076	0.0060	0.0042
		(1.92)**	(1.99)**	(1.93)**	(2.27)**
		[1.99]**	[2.14]**	[1.92]**	[1.99]**
IFRS SALES%	+	−0.0009	−0.0005	−0.0014	−0.0014
		(−0.22)	(−0.12)	(−0.34)	(−0.33)
		[−0.18]	[−0.06]	[−0.30]	[−0.19]
LIFO	−	−0.0017	−0.0013	−0.0019	−0.0025
		(−0.67)	(−0.56)	(−0.71)	(−0.92)
		[−0.84]	[−0.73]	[−0.90]	[−1.00]
SIZE	?	0.0018	0.0014	0.0016	0.0017
		(1.93)*	(1.74)	(1.88)*	(1.87)*
		[1.33]	[1.01]	[1.25]	[1.22]
HERFINDAHL	?	−0.0003	0.0023	−0.0020	−0.0018
		(−0.05)	(0.34)	(−0.23)	(−0.28)
		[−0.21]	[0.22]	[−0.21]	[−0.37]
LEVERAGE	?	0.0050	0.0047	0.0049	0.0049
		(1.15)	(1.19)	(1.23)	(1.12)
		[1.67]	[1.74]	[1.78]*	[1.64]
TURNOVER	?	0.0027	0.0025	0.0024	0.0024
		(1.78)*	(1.65)	(1.48)	(1.56)
		[2.72]**	[2.59]**	[2.37]**	[2.39]**
BIG4	?	0.0019	0.0014	0.0015	0.0023
		(2.52)**	(2.17)**	(1.97)*	(2.72)**
		[3.26]***	[2.98]***	[2.62]**	[3.29]***
n		59,285	59,284	56,254	51,501
R^2		0.0079	0.0090	0.0085	0.0077

(continued on next page)

TABLE 4 (continued)

Panel B: Correlations *D(IFRS)* Measures

	D(IFRS)	*D(IFRS SIMILAR)*	*D(IFRS SIC3)*
D(IFRS SIMILAR)	0.608		
D(IFRS SIC3)	0.344	0.542	
D(IFRS ICB)	0.514	0.286	0.328

*, **, *** Indicate significance at 0.10, 0.05, and 0.01 levels, respectively (two-sided, unless direction is predicted).
[a] *CONVERGENCE* is defined as follows:
(1) *D(IFRS)*
(2) *D(IFRS SIMILAR)*
(3) *D(IFRS SIC3)*
(4) *D(IFRS ICB)*
Table 4, Panel A presents main cross-sectional analyses. Each model includes a different measure of convergence benefits; numbers correspond to variables defined above. t-statistics in parentheses are based on White standard errors that are also clustered at event level. t-statistics in square brackets are from comparison of coefficients for three-day event-returns (reported in table) and coefficients with three-day non-event returns as dependent variable. The non-event coefficient is used as the benchmark value instead of assuming $H_0 = 0$. Spearman correlations between the different convergence benefit measures are provided in Panel B; all correlations are significant at less than the 0.01 level. Variables are as defined in Appendix B.

Third, the significant and positive coefficient for *D(IFRS)* is consistent with investors expecting IFRS adoption to result in convergence benefits for firms in industries where IFRS is already widely adopted. This finding supports the SEC's claim that the benefits of IFRS adoption are likely to be most pronounced for firms in IFRS-predominant industries.[18] In Section VI, we conduct additional analyses to gain more insight into the nature of these convergence benefits.

Fourth, the findings on the cost impact variables are inconsistent with the theoretical predictions. The coefficient for *IFRS SALES%* is insignificant, which does not support the idea that firms with sales in IFRS countries would benefit significantly from reduced preparation costs. One explanation could be that from the investors' point of view, the cost impact is not important enough to lead to a significant response to news about IFRS adoption. Another explanation is that the tests lack power. *IFRS SALES%* captures the firms that operate in IFRS countries, but not necessarily those that are required to use IFRS. Unfortunately, we cannot identify which U.S. firms are legally required to report in IFRS for their foreign subsidiaries.

Also, firms that apply LIFO do not react more negatively to IFRS adoption. Although the *LIFO* variable has the expected negative sign, it is not significant ($p = 0.51$). Since this variable does not take into account the extent to which LIFO is used, and thus what the cost impact would be of adopting IFRS, we also use the ratio LIFO Reserve/Total Assets and identify firms that use LIFO as their primary inventory-valuation method as alternative proxies, but obtain similar results. Similar to the explanations for the insignificance of *IFRS SALES%*, it could be that investors do not expect that disallowing the use of LIFO under IFRS will have a major cost impact and therefore do not react more negatively to IFRS adoption events.

Fifth, the market reaction is positively and significantly related to the control variables *SIZE*, *TURNOVER*, and *BIG4*. As explained in Section IV, one explanation for the positive relation

[18] We also examine whether event returns are positively related to the proportion of global industry peers using IFRS. Untabulated results indicate that the market reaction is significantly positively associated with this proportion, similar to the results for *D(IFRS)*. This is consistent with investors valuing convergence benefits, and with these benefits increasing with the number of firms that use IFRS in a given industry.

between size and event returns is that convergence benefits could be larger, since larger firms are more likely to compete and be compared on a global basis. Moreover, larger and more liquid firms attract institutional ownership and analyst following (O'Brien and Bhushan 1990; Gompers and Metrick 2001). Since institutional investors and analysts have been shown to favor conformity in accounting choices or outcomes (Bradshaw et al. 2004; De Franco et al. 2011), this might be another reason for the positive coefficients for these variables. Finally, the positive coefficient for *BIG4* indicates that investors react more favorably to IFRS adoption for firms with a Big 4 auditor, consistent with these auditors being more able to support the transition from U.S. GAAP to IFRS.

VI. SENSITIVITY ANALYSES

In Section V, we find that the coefficient for *D(IFRS)* is positive and significant, suggesting that investors expect IFRS to lead to convergence benefits. We examine whether this result is robust to different definitions and refinements of our measure of convergence benefits. We also conduct additional analyses to check the sensitivity of our findings to potential confounding events, which is particularly important given that our events take place during the recent financial crisis.

Further Evidence on Convergence Benefits

We perform several analyses to examine whether *D(IFRS)* accurately captures net convergence benefits for U.S. firms. First, as explained earlier, we expect net convergence benefits to be most pronounced for U.S. firms in industries where IFRS is most commonly used by industry peers, assuming that firms within an industry are economically similar. Since there is variation in the extent to which firms are similar within an industry, we refine *D(IFRS)* by explicitly incorporating the degree of economic similarity. We calculate the absolute correlation in daily stock returns between the U.S. firm and each of the top 20 peers using all trading days in the calendar year preceding an event, resulting in 20 correlations per firm. A high absolute correlation indicates that firms are affected to a similar extent by the same economic events, suggesting a high degree of economic similarity.[19] We calculate the average of these 20 correlations for each U.S. firm, and if this average correlation is higher than the median value across all U.S. firms in the same year, we label this firm as having a high degree of economic similarity with the top 20 peers. The variable *D(IFRS SIMILAR)* is equal to 1 for a particular firm if most of its top 20 peers use IFRS and its stock returns are highly correlated with those of the peers. We replace *D(IFRS)* with *D(IFRS SIMILAR)* in the regression analysis; the results are presented in the second column of Table 4, Panel A. The coefficient for *D(IFRS SIMILAR)* is positive and has a slightly higher statistical significance than for *D(IFRS)*.

In a related untabulated analysis, we examine whether investors expect IFRS to make dissimilarities between dissimilar firms more apparent. We calculate the average absolute correlation of a firm with the top 20 largest peers in all *SIC2* industries except for its own, and we identify the industry that has the lowest average correlation. *D(IFRS DISSIMILAR)* equals 1 if IFRS is the predominant standard in that industry. If investors expect IFRS to significantly increase this

[19] We use absolute correlations because we focus on the *magnitude* of the impact of economic events on firms rather than the *direction*. General economic events, such as changes in oil prices, may affect similar firms to the same extent and cause stock returns to move in the same direction. However, if one firm for instance announces increased R&D spending, investors may perceive this as good news for this firm but as bad news for an economically similar competing firm. Thus, we would observe a positive reaction for the announcing firm, but a negative reaction for the competing firm. In such cases, a high negative correlation may also indicate a high degree of economic similarity, which is why we focus on high absolute correlations as a measure of similarity. However, empirically there are few instances where there is a high negative absolute correlation between firms, and the findings are similar if we use non-absolute correlations to capture similarity.

aspect of comparability, then *D(IFRS DISSIMILAR)* should have a positive significant coefficient. However, if investors expect no comparability increase after IFRS, then *D(IFRS DISSIMILAR)* would have an insignificant or significantly negative coefficient. We find that *D(IFRS DISSIMILAR)* has a positive but insignificant coefficient in the regressions, suggesting that investors do not expect IFRS to make the dissimilarities between firms more apparent. However, our measures of convergence benefits remain significant.

Second, to assess the sensitivity of our convergence result to industry definitions, we define *D(IFRS SIMILAR)* at the three-digit SIC level and use this variable, labeled *D(IFRS SIC3)*, instead of *D(IFRS)*. We also define *D(IFRS)* using ICB subsector codes. We find similar results using these two variables, as shown in the last two columns of Table 4, Panel A. Overall, the tests discussed above show that the presence of convergence benefits is a plausible explanation for the higher market reaction in IFRS-predominant industries. The correlations between the different convergence benefit measures are also provided in Panel B of Table 4; all are positive and significant as expected, and most of them are reasonably high (above 0.5).

Next, we examine whether the coefficient on *D(IFRS)* reflects investors' expectations of higher reporting quality under IFRS rather than net convergence benefits. We test whether reporting quality is systematically lower in IFRS-predominant industries using measures of earnings quality, following Francis et al. (2004). If this is the case, then an alternative explanation for the *D(IFRS)* coefficient is that investors expect IFRS to improve reporting quality for firms in those industries. We also examine the sensitivity of the findings in Table 4 and *D(IFRS)* in particular to inclusion of these earnings quality measures in the analyses. Table 5 presents our analyses using seven commonly used earnings quality proxies (accrual quality, earnings persistence, predictability and smoothness, value relevance, earnings timeliness, and conservatism) following Francis et al. (2004), measured in the year preceding the events. First, as Panel A shows, we find that most of these measures (accrual quality, persistence, smoothness, conservatism, and timeliness) indicate that the quality of reporting is *not* lower in IFRS-predominant industries, which is inconsistent with reporting quality being a larger concern for firms in these industries. Second, we find that the coefficients on the convergence benefit proxies are still significant when we include the earnings quality measures in the regressions, although *D(IFRS)* is only weakly significant. For parsimony, we only report the analyses including the accrual quality proxies in Panel B of Table 5, but the findings are similar if we use any of the other earnings quality measures. Overall, these results suggest that *D(IFRS)* proxies for the extent of net convergence benefits rather than expectations of increased reporting quality that could result from IFRS adoption.

Credible Implementation of IFRS

Convergence might lower the costs of comparing information for investors, but this is likely to occur only if firms implement IFRS properly. The intuition is similar to that in DeFond et al. (2011), who use the change in mutual fund ownership after IFRS adoption to infer whether IFRS led to higher reporting comparability, which is a potential benefit of convergence. They find that mutual fund ownership only increases in the event of a "credible increase in uniformity," i.e., if the implementation of IFRS is well enforced. Similarly, we use the World Bank rule-of-law scores from Kaufmann et al. (2009) to measure the implementation quality of IFRS. We assign these scores to each firm in the top 20 peer group based on its country of origin and calculate the average rule-of-law score for each group. We define two variables, $D(IFRS_{Weak})$ and $D(IFRS_{Strong})$, to distinguish between industries where IFRS is commonly used but with different levels of implementation quality. $D(IFRS_{Weak})$ (respectively $D(IFRS_{Strong})$) is equal to 1 if IFRS is commonly used and the average rule-of-law score for the peer group is below (above) the median value across industries. If implementation quality matters, then the coefficient for $D(IFRS_{Strong})$ should be higher than that for

TABLE 5
Reporting Quality or Convergence Benefits

Panel A: Average Earnings Quality by Industry Type (*D(IFRS)* = 0/1)

Earnings Quality Measure[a]	A *D(IFRS)* = 0	B *D(IFRS)* = 1	Difference (B − A)
ACCRUAL QUALITY	0.0498	0.0265	−0.0233***
PERSISTENCE	−0.3758	−0.4034	−0.0276***
PREDICTABILITY	0.9934	1.1339	0.1405***
SMOOTHNESS	1.0697	0.9136	−0.1561***
VALUE RELEVANCE	−0.3879	−0.3645	0.0234***
TIMELINESS	−0.4496	−0.4447	0.0049
CONSERVATISM	−0.0109	−0.5293	−0.5184*

[a] Note that *lower* values correspond to *higher* earnings quality (following Francis et al. [2004]). Hence a negative difference in the last column indicates that the earnings quality in industries where *D(IFRS)* = 1 is *higher* than in industries where *D(IFRS)* = 0.

Panel B: Cross-Sectional Analyses Including Accrual Quality

$$CR_{STOXX\ ex\ Am.} = f(EXTR,\ INSUR,\ HI\text{-}LIT,\ CONVERGENCE,\ IFRS\ SALES\%,\ LIFO,\ ACCRUAL\ QUALITY,\ Control\ Variables)^{b}$$

	Prediction	(1)	(2)	(3)	(4)
Intercept	?	−0.0188	−0.0169	−0.0173	−0.0179
		(−2.42)**	(−2.29)**	(−2.33)**	(−2.23)**
		[−2.20]**	[−2.03]*	[−2.11]**	[−1.94]*
EXTR	−	0.0112	0.0115	0.0114	0.0116
		(1.67)	(1.72)	(1.68)	(1.69)
		[1.61]	[1.68]	[1.61]	[1.59]
INSUR	−	0.0100	0.0092	0.0091	0.0111
		(1.35)	(1.28)	(1.27)	(1.46)
		[1.36]	[1.27]	[1.28]	[1.48]
HI-LIT	−	−0.0023	−0.0021	−0.0022	−0.0022
		(−1.39)*	(−1.35)*	(−1.44)*	(−1.29)
		[−1.20]	[−1.10]	[−1.26]	[−1.21]
CONVERGENCE[b]	+	0.0023	0.0044	0.0044	0.0047
		(1.42)*	(1.94)**	(1.72)*	(2.47)**
		[1.47]*	[2.18]**	[1.69]*	[2.23]**
IFRS SALES%	+	−0.0001	0.0000	0.0002	−0.0017
		(−0.02)	(0.01)	(0.04)	(−0.44)
		[0.10]	[0.17]	[0.19]	[−0.26]
LIFO	−	0.0007	0.0009	0.0010	0.0000
		(0.50)	(0.64)	(0.73)	(−0.02)
		[0.51]	[0.68]	[0.62]	[0.10]
ACCRUAL QUALITY	?	−0.0083	−0.0080	−0.0072	0.0031
		(−0.43)	(−0.42)	(−0.37)	(0.14)
		[0.38]	[0.40]	[0.39]	[0.84]
SIZE	?	0.0013	0.0011	0.0012	0.0012
		(1.73)	(1.61)	(1.65)	(1.58)
		[1.23]	[1.06]	[1.14]	[1.03]

(continued on next page)

TABLE 5 (continued)

	Prediction	(1)	(2)	(3)	(4)
HERFINDAHL	?	−0.0041	−0.0026	−0.0015	−0.0041
		(−0.42)	(−0.26)	(−0.16)	(−0.39)
		[−0.39]	[−0.21]	[0.17]	[−0.38]
LEVERAGE	?	0.0039	0.0037	0.0035	0.0044
		(0.99)	(0.93)	(0.91)	(1.08)
		[1.89]*	[1.81]*	[1.79]*	[1.91]*
TURNOVER	?	0.0019	0.0018	0.0018	0.0018
		(1.09)	(1.00)	(1.02)	(1.02)
		[1.77]*	[1.69]	[1.72]	[1.67]
BIG4	?	0.0021	0.0020	0.0020	0.0020
		(2.19)**	(2.16)**	(2.18)**	(2.16)**
		[2.98]***	[2.97]**	[2.95]**	[3.15]***
n		28,117	28,117	28,117	25,378
R^2		0.0087	0.0091	0.0091	0.0089

*, **, *** Indicate significance at 0.10, 0.05, and 0.01 levels, respectively (two-sided, unless direction is predicted).
[b] *CONVERGENCE* is defined as follows:
(1) *D(IFRS)*
(2) *D(IFRS SIMILAR)*
(3) *D(IFRS SIC3)*
(4) *D(IFRS ICB)*
Table 5 presents sensitivity analyses to assess whether reporting quality is a credible explanation for convergence benefit proxies. Panel A presents average earnings quality in industries that are IFRS-predominant versus those that are not. If reporting quality explains the positive coefficient on the convergence benefit proxies in Table 4, we expect to observe *lower* reporting quality (i.e., *high* values on these earnings quality measures) for industries where *D(IFRS)* = 1. Panel B presents the results for cross-sectional analyses controlling for *ACCRUAL QUALITY* (for parsimony, results including the other earnings quality measures are untabulated). Calculation of earnings quality measures follows Francis et al. (2004), as these are estimated on a ten-year rolling window basis, the number of observations drops compared to Table 4. For calculation of t-statistics (in parentheses), see footnote to Table 4. Variables are as defined in Appendix B.

$D(IFRS_{Weak})$. We find that $D(IFRS_{Strong})$ is indeed positively and significantly associated with the market reaction, whereas $D(IFRS_{Weak})$ is insignificant. However, the difference in the two coefficients is not statistically significant ($p = 0.74$). A possible explanation is the fact that convergence benefits encompass more than reduced costs of comparing information. For example, enhanced consistency in global accounting practices applied by preparers and their auditors is a convergence benefit that does not rely on the implementation quality of peers. Hence, since both $D(IFRS_{Weak})$ and $D(IFRS_{Strong})$ capture net convergence benefits to some extent, it is plausible that the coefficients do not differ significantly.

Foreign Institutional Ownership and Analyst Following

Prior research has shown that there is a demand for conformity in reporting choices across firms from analysts and institutional investors (e.g., Bradshaw et al. 2004; De Franco et al. 2011; DeFond et al. 2011). We therefore investigate whether the market reactions are higher when there is greater demand for convergence and include analyst following and a variable for high foreign institutional ownership in the regression model.[20] We find that analyst following is not

[20] For parsimony, we have tabulated the sensitivity analyses with *D(IFRS SIMILAR)*, but we find similar results when we measure convergence benefits with *D(IFRS)*, *D(IFRS SIC3)*, and *D(IFRS ICB)*, unless stated otherwise. Results are available from the authors upon request.

significantly related to the market reaction, but if a firm has a higher number of foreign institutional owners, the market reaction is significantly higher. Table 6 also shows that despite the inclusion of analyst following and institutional ownership, *D(IFRS SIMILAR)* remains significantly and positively related to the market reaction. This evidence supports the main finding that investors appear to value convergence benefits of IFRS adoption within IFRS-predominant industries.

Other News and Market Adjustment

We take several steps to ensure our results are not sensitive to the chosen market adjustment, other news events, and selection of events.

First, we assess the sensitivity of our results to an alternative market adjustment by subtracting the STOXX 600 Asia-Pacific index return instead of the STOXX 1800 Global ex America index return. The results for the overall market reaction (Table 3) are similar for both adjustments. We also find that the cross-sectional analyses are generally robust to this alternative adjustment (column 3 of Table 6 shows the regression results for the *D(IFRS SIMILAR)* specification): the coefficients for the convergence benefit proxies are generally similar to those in the main analysis, except for *D(IFRS)*, which loses significance.

Second, Tables 4 to 6 provide additional information on whether crisis-related news is driving our results. The tables report alternative t-statistics (in brackets), which are based on a comparison of the coefficients in the main analysis with coefficients from the same models when we replace the event returns with non-event returns (see also Armstrong et al. 2010). We do this to eliminate the possibility that we are merely capturing systematic relations between returns and firm characteristics. We find that most of the results are unaltered and that the significance of our metrics of convergence benefits is enhanced by this procedure.

Third, we search in Factiva for other major news in our sample period (April 2007 to January 2009). News about the problems with mortgage securities and the economic downturn was reported in several event windows. This may explain why some of our returns in Table 3 are inconsistent with their predicted signs. For example, for the events on October 24, 2007 and November 7 and 15, 2007, the *Wall Street Journal* reported news on disappointing economic statistics, downgrades of mortgage securities, and slowing economic growth and inflation. The most strongly affected event is October 13, 2008, which overlaps with the 2008 stock market crash, as mentioned earlier. We also repeat the cross-sectional analyses in Table 4 excluding this event; the last column of Table 6 shows the results for *D(IFRS SIMILAR)*. Although the cross-sectional results are weaker when we exclude this event, the results for any of the convergence benefit proxies do not seem to be driven by any particular event. We find similar results if we exclude any other event from the sample. Also, unreported analyses show that excluding banks from the sample (firms with *SIC2* equal to 60 or 61), which were most severely affected by the crisis, does not affect the tenor of our results.

We acknowledge that it is impossible to perfectly adjust for potentially confounding effects, since major news events can occur every day. However, the inferences from our main cross-sectional analyses are similar when we control for confounding events, and we believe that they provide insight into which firms investors expect will benefit from IFRS adoption.

VII. CONCLUSION

The goal of this study is to provide empirical evidence for the costs and benefits of IFRS adoption by U.S. firms from the investors' point of view. We use the stock market reaction to events that affect the likelihood of IFRS adoption and examine whether this varies cross-sectionally in a predictable manner. We find that investors react more positively to events that increase the

TABLE 6

Additional Sensitivity Analyses

	Prediction	(1) Implementation Quality IFRS	(2) Analyst Following/ Foreign Institutional Ownership	(3) STOXX 600 AP Market Adjustment	(4) Exclude Event 13
Intercept	?	−0.0262	−0.0151	−0.0107	−0.0172
		(−2.63)**	(−1.29)	(−0.92)	(−2.12)*
		[−2.04]*	[−0.69]	[−0.01]	[−1.35]
EXTR	–	0.0112	0.0119	0.0119	0.0076
		(1.51)	(1.55)	(1.58)	(1.13)
		[1.72]	[1.76]	[1.84]	[1.38]
INSUR	–	0.0052	0.0037	0.0039	0.0001
		(1.00)	(0.80)	(0.87)	(0.03)
		[0.86]	[0.61]	[0.65]	[−0.27]
HI-LIT	–	−0.0024	−0.0020	−0.0024	−0.0018
		(−2.18)**	(−1.93)**	(−2.26)**	(−1.63)*
		[−1.74]*	[−1.50]*	[−1.68]*	[−1.13]
$D(IFRS_{Weak})$	?	0.0026			
		(1.14)			
		[1.12]			
$D(IFRS_{Strong})$	+	0.0037			
		(1.69)*			
		[1.82]**			
D(IFRS SIMILAR)	+		0.0073	0.0071	0.0052
			(2.02)**	(1.79)**	(1.60)*
			[2.22]**	[2.04]**	[1.78]**
ANALYST	+		0.0000		
			(0.01)		
			[0.32]		
FORINSTIT	+		0.0035		
			(1.94)**		
			[1.65]*		
IFRS SALES%	+	−0.0011	−0.0007	−0.0007	0.0018
		(−0.25)	(−0.17)	(−0.18)	(0.55)
		[−0.22]	[−0.09]	[−0.01]	[0.62]
LIFO	–	−0.0018	−0.0014	−0.0007	0.0001
		(−0.72)	(−0.58)	(−0.34)	(0.03)
		[−0.91]	[−0.70]	[−0.50]	[−0.17]
SIZE	?	0.0018	0.0009	0.0011	0.0011
		(1.97)*	(0.79)	(1.06)	(1.32)
		[1.37]	[0.20]	[0.41]	[0.59]
HERFINDAHL	?	−0.0005	0.0031	−0.0033	0.0045
		(−0.08)	(0.45)	(−0.48)	(0.68)
		[−0.25]	[0.32]	[−0.45]	[0.56]
LEVERAGE	?	0.0053	0.0046	0.0053	0.0031
		(1.09)	(1.19)	(1.27)	(0.80)
		[1.56]	[1.74]	[1.83]*	[1.35]

(continued on next page)

TABLE 6 (continued)

	Prediction	(1) Implementation Quality IFRS	(2) Analyst Following/ Foreign Institutional Ownership	(3) STOXX 600 AP Market Adjustment	(4) Exclude Event 13
TURNOVER	?	0.0026	0.0017	0.0029	0.0020
		(1.76)*	(1.00)	(1.77)*	(1.25)
		[2.71]**	[1.84]*	[2.74]**	[2.14]**
BIG4	?	0.0018	0.0011	0.0020	0.0013
		(2.60)**	(1.65)	(2.43)**	(1.89)*
		[3.37]***	[2.51]**	[3.14]***	[2.65]**
n		59,285	59,284	59,284	55,784
R^2		0.0079	0.0093	0.0080	0.0058

*, **, *** Indicate significance at 0.10, 0.05, and 0.01 levels, respectively (two-sided, unless direction is predicted).
Table 6 presents additional sensitivity analyses, where we differentiate between industries where IFRS is predominant based on average IFRS implementation quality (column 1), include analyst following and foreign institutional ownership (column 2), use an alternative market index to adjust returns (STOXX 600 Asia-Pacific index) (column 3), and exclude event 13 from the analyses (column 4). Variables are as defined in Appendix B. For calculation of t-statistics (in parentheses and square brackets), see footnote to Table 4.

likelihood of IFRS adoption in cases where IFRS is expected to result in convergence benefits. We find a significantly more positive market reaction for firms operating in industries where IFRS is the predominant choice worldwide, for larger and more liquid firms that are more likely to attract investors who stand to benefit from convergence, and for firms with high foreign institutional ownership. Collectively, these findings suggest that investors expect U.S. adoption of IFRS to result in net convergence benefits. Further, investors in firms with high litigation risk respond less positively to events that increase the likelihood of IFRS, consistent with the notion that IFRS may increase the likelihood of expensive legal challenges or may lead to overly conservative behavior by these firms to avoid litigation.

The findings of this paper must be interpreted carefully in light of several limitations. First, our focus is only on investors, rather than on all parties that could be affected by the change. Second, the methodology relies on a correct identification of events and requires that event-related information be incorporated into stock prices rapidly and without bias (Armstrong et al. 2010). We have carefully identified the relevant events and dates, but we cannot exclude the possibility that participants were privy to relevant information prior to the dates identified here. Third, the findings relate to the expected effects of IFRS adoption, rather than to the realized effects, and should therefore be seen as preliminary evidence for the effects of IFRS adoption.

Despite these limitations, our findings are relevant to the current debate on whether the U.S. should switch to IFRS. The SEC has stated that the transition should be made only if it benefits U.S. investors and capital markets, and this paper provides evidence relevant to that issue. A final contribution is our finding that despite the ongoing efforts of the IASB and FASB to reduce the differences between IFRS and U.S. GAAP, investors still react positively to news that increases the likelihood of having a single set of standards. Our findings highlight the importance of convergence benefits to investors and show that there are both costs and benefits to the use of a common global accounting standard.

REFERENCES

Armstrong, C. S., M. E. Barth, A. D. Jagolinzer, and E. J. Riedl. 2010. Market reaction to the adoption of IFRS in Europe. *The Accounting Review* 85 (1): 31–61.

Ashbaugh, H., and M. Pincus. 2001. Domestic accounting standards, international accounting standards, and the predictability of earnings. *Journal of Accounting Research* 39 (3): 417–434.

Ball, R., A. Robin, and J. S. Wu. 2003. Incentives versus standards: Properties of accounting income in four East Asian countries. *Journal of Accounting and Economics* 36 (1-3): 235–270.

Ball, R., and L. Shivakumar. 2005. Earnings quality in UK private firms: Comparative loss recognition timeliness. *Journal of Accounting and Economics* 39 (1): 83–128.

Ball, R., S. P. Kothari, and A. Robin. 2000. The effect of international institutional factors on properties of accounting earnings. *Journal of Accounting and Economics* 29 (1): 1–51.

Barth, M. E. 2008. Global financial reporting: Implications for U.S. academics. *The Accounting Review* 83 (5): 1159–1179.

Barth, M. E., G. Clinch, and T. Shibano. 1999. International accounting harmonization and global equity markets. *Journal of Accounting and Economics* 26 (1-3): 201–235.

Barth, M. E., W. R. Landsman, and M. H. Lang. 2008. International accounting standards and accounting quality. *Journal of Accounting Research* 46 (3): 467–498.

Beuselinck, C., P. Joos, I. K. Khurana, and S. Van der Meulen. 2011. *Mandatory IFRS Reporting and Stock Price Informativeness*. Working paper, Tilburg University and the University of Missouri at Columbia.

Bothwell, J. L. 2009. *Adopting International Financial Reporting Standards for Use in the United States: An Economic and Public Policy Perspective*. Study commissioned by the FASB and included as part of the FAF response to the SEC roadmap. Available at: http://www.fasb.org/cs/BlobServer?blobkey=id&blobwhere=1175818992147&blobheader=application%2Fpdf&blobcol=urldata&blobtable=MungoBlobs

Botosan, C. A. 1997. Disclosure level and the cost of equity capital. *The Accounting Review* 72 (3): 323–349.

Bradshaw, M. T., B. J. Bushee, and G. S. Miller. 2004. Accounting choice, home bias, and U.S. investment in non-U.S. firms. *Journal of Accounting Research* 42 (5): 795–841.

Burgstahler, D. C., L. Hail, and C. Leuz. 2006. The importance of reporting incentives: Earnings management in European private and public firms. *The Accounting Review* 81 (5): 983–1016.

Chi, S. 2009. *Simultaneous Presence of Different Domestic GAAPs and Investors' Limited Attention Bias in U.S. Equity Markets: Implications for Convergence*. Working paper, 2010 AAA Financial Accounting and Reporting Section.

Christensen, H. B., E. Lee, and M. Walker. 2007. Cross-sectional variation in the economic consequences of international accounting harmonization: The case of mandatory IFRS adoption in the UK. *The International Journal of Accounting* 42 (4): 341–379.

Curran, R. 2008. Dow Jones Industrials lost 18% in their worst week ever, Wachovia gains 43%, but lost on week; Look at me; It's a flush; It's a panic. *Wall Street Journal* (October 11): B3.

Daske, H., L. Hail, C. Leuz, and R. Verdi. 2008. Mandatory IFRS reporting around the world: Early evidence on the economic consequences. *Journal of Accounting Research* 46 (5): 1085–1142.

De Franco, G., S. P. Kothari, and R. Verdi. 2011. The benefits of financial statement comparability. *Journal of Accounting and Economics* 49 (4): 895–931.

Dechow, P. M., and I. D. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (Supplement): 35–59.

DeFond, M. L., X. Hu, M. Hung, and S. Li. 2011. The impact of mandatory IFRS adoption on foreign mutual fund ownership: The role of comparability. *Journal of Accounting and Economics* 51 (3): 240–258.

Diamond, D. W., and R. E. Verrecchia. 1991. Disclosure, liquidity, and the cost of capital. *The Journal of Finance* 46 (4): 1325–1359.

Drake, M. S., L. A. Myers, and L. Yao. 2010. *Are Liquidity Improvements Around the Mandatory Adoption of IFRS Attributable to Comparability Effects or to Quality Effects?* Working paper, 2010 AAA Financial Accounting and Reporting Section.

Fama, E. F., and K. R. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43: 153–193.

Field, L., M. Lowry, and S. Z. Shu. 2005. Does disclosure deter or trigger litigation? *Journal of Accounting and Economics* 39: 487–507.

Financial Accounting Foundation (FAF). 2009. *Response letter to Roadmap for the Potential Use of Financial Statement Prepared in Accordance with International Financial Reporting Standards (IFRS) By U.S. Issuers*. Available at: http://www.sec.gov/comments/s7-27-08/s72708-65.pdf

Financial Accounting Standards Board (FASB). 1980. *Qualitative Characteristics of Accounting Information*. Statement of Financial Accounting Concepts No. 2. Available at: http://www.fasb.org/pdf/con2.pdf

Financial Accounting Standards Board (FASB). 2008. *Completing the February 2006 Memorandum of Understanding: A Progress Report and Timetable for Completion*. Available at: http://www.fasb.org/intl/mou_09-11-08.pdf

Financial Crisis Advisory Group (FCAG). 2009. *Report of the Financial Crisis Advisory Group*. Available at: http://www.ifrs.org/News/Press-Releases/Documents/FCAGReportJuly2009.pdf

Financial Executives International (FEI). 2004. *FEI Special Survey on Sarbanes-Oxley Section 404 Implementation Executive Summary*. Available at: http://www.404institute.com/docs/SOXSurveyJuly.

Francis, J., R. LaFond, P. M. Olsson, and K. Schipper. 2004. Costs of equity and earnings attributes. *The Accounting Review* 79 (4): 967–1010.

Gompers, P. A., and A. Metrick. 2001. Institutional investors and equity prices. *Quarterly Journal of Economics* 116: 229–259.

Hail, L., C. Leuz, and P. D. Wysocki. 2010. Global accounting convergence and the potential adoption of IFRS by the U.S. (part I): Conceptual underpinnings and economic analysis. *Accounting Horizons* 24 (3): 355–394.

Hand, J. R. M. 2003. The increasing returns-to-scale of intangibles. In *Intangible Assets: Values, Measures and Risks*, edited by B. Lev. New York: Oxford University Press.

Hewitt, C. W. 2008. Speech by SEC staff. Remarks before the 2008 AICPA National Conference on Current SEC and PCAOB Developments. Available at: http://www.sec.gov/news/speech/2008/spch120808cwh.htm

Hung, M. 2000. Accounting standards and value relevance of financial statements: An international analysis. *Journal of Accounting and Economics* 30 (3): 401–420.

International Accounting Standards Board (IASB). 2011. *Who We Are and What We Do*. Available at: http://www.ifrs.org/The-organisation/Documents/WhoWeAre2012MarchEnglish.pdf

Johnson, S. 2009. Mary Schapiro vows to be tough enforcer. Available at: http://www3.cfo.com/article.cfm/12958299?currpage=0

Kasznik, R., and B. Lev. 1995. To warn or not to warn: Management disclosures in the face of an earnings surprise. *The Accounting Review* 70 (1): 113–134.

Kaufmann, D., A. Kraay, and M. Mastruzzi. 2009. *Governance Matters VIII: Aggregate and Individual Governance Indicators, 1996–2008*. Working Paper No. 4978, World Bank Policy Research.

Kothari, S. P., K. Ramanna, and D. J. Skinner. 2010. Implications for GAAP from an analysis of positive research in accounting. *Journal of Accounting and Economics* 50: 246–286.

Kroeker, J. 2009. Speech by SEC staff: Remarks before the 2009 AICPA National Conference on Current SEC and PCAOB Developments. Available at: http://www.sec.gov/news/speech/2009/spch120709jlk.htm

Lambert, R., C. Leuz, and R. E. Verrecchia. 2007. Accounting information, disclosure and the cost of capital. *Journal of Accounting Research* 45 (2): 385–420.

Lang, M., M. Maffett, and E. Owens. 2010. *Earnings Co-Movement and Accounting Comparability: The Effects of Mandatory IFRS Adoption*. Working paper, The University of North Carolina.

Leuz, C. 2007. Was the Sarbanes–Oxley Act of 2002 really this costly? A discussion of evidence from event returns and going-private decisions. *Journal of Accounting and Economics* 44: 146–165.

Leuz, C., and R. E. Verrecchia. 2000. The economic consequences of increased disclosure. *Journal of Accounting Research* 38 (Supplement): 91–124.

Leuz, C., D. Nanda, and P. D. Wysocki. 2003. Earnings management and investor protection: An international comparison. *Journal of Financial Economics* 69 (3): 505–527.

Li, S. 2010. Does mandatory adoption of International Financial Reporting Standards in the European Union reduce the cost of equity capital? *The Accounting Review* 85 (2): 607–636.

Lin, S., and P. Tanyi. 2010. *Market Reaction to the Potential Adoption of International Accounting Standards in the U.S.* Working paper, Florida International University.

Matsumoto, D. A. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (3): 483–514.

O'Brien, P. C., and R. Bhushan. 1990. Analyst following and institutional ownership. *Journal of Accounting Research* 28 (Supplement): 55–76.

Pacter, P. 2005. What exactly is convergence? *International Journal of Accounting, Auditing and Performance Evaluation* 2 (1): 67–83.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22 (1): 435–480.

Securities and Exchange Commission (SEC). 2007. *Concept Release on Allowing U.S. Issuers to Prepare Financial Statements in Accordance with International Financial Reporting Standards*. Release Nos. 33-8831, 34-56217. Washington, DC: SEC.

Securities and Exchange Commission (SEC). 2008. *Roadmap for the Potential Use of Financial Statements Prepared in Accordance with International Financial Reporting Standards by U.S. Issuers*. Release No. 33-8982. Washington, DC: SEC.

Sletten, E., and K. Ramanna. 2009. *Why Do Countries Adopt International Financial Reporting Standards*? Working paper no. 09-102, Harvard Business School Accounting and Management Unit.

Steverman, B. 2008. Stock market crash: Understanding the panic. *Businessweek* (October 13). Available at: http://www.businessweek.com/investor/content/oct2008/pi2008109_360708.htm

Tweedie, D. 2006. Prepared statement of Sir David Tweedie, Chairman of the International Accounting Standards Board before the Economic and Monetary Affairs Committee of the European Parliament. Available at: http://www.ifrs.org/News/Press-Releases/Documents/ECOFINDavidTweediestatement.pdf

Wu, J. S., and I. X. Zhang. 2010. *Accounting Integration and Comparability: Evidence from Relative Performance Evaluation around IFRS Adoption*. Working paper, University of Rochester and University of Minnesota.

Yip, R. W. Y., and D. Young. 2011. Does Mandatory IFRS Adoption Improve Information Comparability? *The Accounting Review* 87 (5): 1767–1789).

Zeff, S. A. 2007. Some obstacles to global financial reporting comparability and convergence at a high level of quality. *The British Accounting Review* 39: 290–302.

Zhang, I. X. 2007. Economic consequences of the Sarbanes-Oxley Act of 2002. *Journal of Accounting and Economics: Conference Issue on Corporate Governance: Financial Reporting, Internal Control, and Auditing* 44 (1-2): 74–115.

APPENDIX A

SAMPLE SELECTION

		Number of Observations
All U.S. firms not missing return data on one of 15 events		63,597
Less: observations for firms with negative book value of equity	−1,987	61,610
Less: firms missing data for control variables:[a]	−2,325[b]	
SIZE	−1,496	
TURNOVER	−1,482	
LEVERAGE	−75	
Final sample main analysis[c]		59,285

[a] *EXTR*, *INSUR*, *HI-LIT*, and *D(IFRS)* are defined at the industry level, for which there are sufficient data available. *IFRS SALES%*, *LIFO*, and *BIG4* are equal to 0 if Compustat does not report foreign sales, inventory valuation method, or auditor.

[b] The total number of observations that we lose is not equal to the sum of these three individual components, since most observations lacking *SIZE* also lack data for *TURNOVER*.

[c] To calculate *D(IFRS SIMILAR)*, *D(IFRS SIC3)*, and *D(IFRS ICB)*, we require either sufficient return data to calculate economic similarity with peers, or a firm's ICB industry code in Datastream, which are not always available. This reduces the number of observations for the cross-sectional analyses with these variables.

APPENDIX B

DEFINITION OF VARIABLES USED IN CROSS-SECTIONAL ANALYSES

Dependent Variables

CR = three-day cumulative raw return centered on event date for U.S. firm (source: CRSP);

STOXX 1800 ex America = three-day cumulative return for DJ STOXX 1800 Global index excluding American firms (source: Datastream);

STOXX 600 Asia-Pacific = three-day cumulative return for DJ STOXX 600 Asia-Pacific index three-day cumulative return (source: Datastream);

$CR_{STOXX\ ex\ Am.}$ = *CR* adjusted for STOXX 1800 ex America (sources: CRSP and Datastream); and

$CR_{STOXX\ A.P.}$ = *CR* adjusted for STOXX 600 Asia-Pacific (sources: CRSP and Datastream).

Test Variables

EXTR = indicator variable equal to 1 if firm has *SIC2* = 13 or 29 (extractive industries), 0 otherwise (source: Compustat);

INSUR = indicator variable equal to 1 if firm has *SIC2* = 63 or 64 (insurance industries), 0 otherwise (source: Compustat);

HI-LIT = indicator variable equal to 1 if firm has SIC of 2833–2836 (pharmaceutical), 3570–3577 (computer hardware), 3600–3674 (electronics), 5200–5961 (retail), 7371–7379 (computer software), or 8731–8734 (R&D), 0 otherwise (source: Compustat);

D(IFRS) = indicator variable equal to 1 if IFRS is the most commonly used set of standards among 20 largest firms in a given industry globally (based on *SIC2*), 0 otherwise. For example, if, for a given industry, ten firms use IFRS, five firms use U.S. GAAP, and the remaining five all use different local GAAPs, *D(IFRS)* = 1 for the firms in this industry. Another example would be that eight firms use IFRS, three use U.S. GAAP, and the other

nine each use a different local GAAP; this industry would also be classified as IFRS-predominant (source: Worldscope);

D(IFRS SIMILAR) = indicator variable equal to 1 if IFRS is the most commonly used set of standards among 20 largest firms in a given industry globally (based on *SIC2*) and average absolute yearly correlation in daily stock returns with these firms exceeds median average value, 0 otherwise. We use daily returns for all trading days in the calendar year preceding an event to calculate the correlation. For a given U.S. firm, the absolute correlation is calculated with each of the 20 largest firm peers separately and then averaged (sources: Worldscope and Datastream);

D(IFRS SIC3) = indicator variable equal to 1 if IFRS is the most commonly used set of standards among 20 largest firms in a given industry globally (based on *SIC3*) and average absolute correlation in daily stock returns with these firms exceeds median in year preceding event, 0 otherwise. Calculation of the correlation is as described for *D(IFRS SIMILAR)* (sources: Worldscope and Datastream);

D(IFRS ICB) = indicator variable equal to 1 if IFRS is the most commonly used set of standards among 20 largest firms in a given industry globally (based on ICB subsectors), 0 otherwise (source: Worldscope);

D(IFRS SALES) = indicator variable equal to 1 if firm has non-zero foreign sales in countries that require IFRS reporting, 0 otherwise (source: Compustat segment file for foreign sales; IASPlus and Sletten and Ramanna [2009] used to identify IFRS countries);

IFRS SALES% = firm's sales in countries that require IFRS relative to its total sales (source: Compustat segment file for foreign sales; IASPlus.com and Sletten and Ramanna [2009] used to identify IFRS countries);

LIFO = indicator variable equal to 1 if firm uses LIFO to value its inventory, 0 otherwise (source: Compustat);

$D(IFRS_{Weak})$ = indicator variable equal to 1 if IFRS is the most commonly used set of standards among the 20 largest firms in a given industry globally (based on *SIC2*) and the average implementation quality of IFRS is below the median value. Implementation quality is measured using the Kaufmann et al. (2009) country-level rule-of-law score. We calculate the average rule-of-law score across firms applying IFRS in industries where IFRS is most commonly used (i.e., *D(IFRS)* = 1) (sources: Worldscope for firms' accounting standards; World Bank and Kaufmann et al. [2009] for rule-of-law scores);

$D(IFRS_{Strong})$ = indicator variable equal to 1 if IFRS is the most commonly used set of standards among 20 largest firms in a given industry globally (based on *SIC2*) and the average implementation quality of IFRS is above the median value. See also $D(IFRS_{Weak})$ for further details on measure of implementation quality (sources: Worldscope for firms' accounting standards; World Bank and Kaufmann et al. [2009] for rule-of-law scores);

ANALYST = number of estimates (*NUMEST*) from I/B/E/S, measured at end of quarter preceding event (source: I/B/E/S); and

FORINSTIT = indicator variable equal to 1 if number of foreign institutional owners exceeds median, measured at end of quarter preceding event (source: Thomson Research).

Earnings Quality Variables (following Francis et al. [2004])

ACCRUAL QUALITY = standard deviation of residuals of the Dechow and Dichev (2002) regression model, estimated over ten-year windows for each firm separately. Lower values of standard deviation indicate higher accrual quality. Dechow and Dichev (2002) model:

$$\frac{ACC_{i,t}}{Assets_{i,t}} = \beta_{0,i} + \beta_{1,i}\frac{CFO_{i,t-1}}{Assets_{i,t}} + \beta_{2,i}\frac{CFO_{i,t}}{Assets_{i,t}} + \beta_{3,i}\frac{CFO_{i,t+1}}{Assets_{i,t}} + \varepsilon_{i,t},$$

where:

$ACC_{i,t}$ = total current accruals of firm i in year t, calculated as the difference between income before extraordinary items and cash flow from operations;
$Assets_{i,t}$ = average assets of firm i in year t;
$CFO_{i,t}$ = cash flow from operations of firm i in year t (source: Compustat);
PERSISTENCE = negative of *AR* coefficient ($\gamma_{1,i}$) in the following first-order autoregressive (*AR1*) model of earnings, estimated over ten-year windows for each firm separately (source: Compustat):

$$EPS_{i,t} = \gamma_{0,i} + \gamma_{1,i}EPS_{i,t-1} + \omega_{i,t};$$

PREDICTABILITY = square root of the error variance from the *AR1* model above (see *PERSISTENCE*) (source: Compustat);
SMOOTHNESS = standard deviation of firm's income before extraordinary items/standard deviation of cash flow from operations, calculated over a ten-year window (source: Compustat);
VALUE RELEVANCE = R^2 of following regression, multiplied by −1 (estimated over ten-year windows for each firm separately):

$$R_{i,t} = \delta_{0,i} + \delta_{1,i}E_{i,t} + \delta_{2,i}\Delta E_{i,t} + \vartheta_{i,t},$$

where:

$R_{i,t}$ = 15-month return of firm i in year t, ending three months after the fiscal year end;
$E_{i,t}$ = income before extraordinary items of firm i in year t, divided by market value of equity at the end of year t–1;
$\Delta E_{i,t}$ = change of firm i's income before extraordinary items in year t, divided by market value of equity at the end of year t–1 (sources: Compustat and CRSP);
TIMELINESS = R^2 of the following regression, multiplied by −1 (estimated over ten-year windows for each firm separately):

$$E_{i,t} = \theta_{0,i} + \theta_{1,i}NEG_{i,t} + \theta_{2,i}R_{i,t} + \theta_{3,i}NEG_{i,t}{*}R_{i,t} + \varphi_{i,t};$$

where:

$NEG_{i,t}$ = 1 if $R_{i,t} < 0$, and 0 otherwise, and other variables as defined above (sources: Compustat and CRSP); and
CONSERVATISM = $-(\theta_{2,i} + \theta_{3,i})/\theta_{2,i}$, see regression model above (sources: Compustat and CRSP).

Control Variables

SIZE = log of firm's market value at end of prior calendar year (source: CRSP);
HERFINDAHL = measure of industry concentration (Herfindahl index): sum of each firm's squared percentage market share, calculated at *SIC2* level (source: Compustat);
LEVERAGE = ratio of firm's total liabilities to total assets (source: Compustat);
TURNOVER = indicator variable equal to 1 if firm's mean daily percentage shares traded during the calendar year is above median for all firms, 0 otherwise (source: CRSP); and

BIG4 = indicator variable equal to 1 if firm's auditor belongs to Big 4, 0 otherwise (source: Compustat).

Non-Event Return Adjustments

We use three-day non-event market-adjusted returns as a benchmark for assessing the significance of the overall market reaction in Table 3 and for the alternative t-statistics of the cross-sectional analyses in Tables 4 to 8 (see Armstrong et al. 2010). We start from the first trading day in 2007, and calculate the cumulative return for three consecutive non-event trading days for each U.S. firm in our sample. We subtract the contemporaneous three-day STOXX index return for the same three days, similar to how we calculate $CR_{STOXX\ ex\ Am.}$ and $CR_{STOXX\ A.P.}$ We repeat this for the next three consecutive non-event trading days (i.e., the non-event three-day windows do not overlap) and for all non-event trading days in 2007 and 2008.

All test, earnings quality, and control variables are measured at the end of the fiscal year preceding an event, unless specified otherwise.

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 611–639

American Accounting Association
DOI: 10.2308/accr-50313

Are Short Sellers Informed? Evidence from the Bond Market

Ambrus Kecskés
Sattar A. Mansi
Virginia Polytechnic Institute and State University

Andrew (Jianzhong) Zhang
University of Nevada, Las Vegas

ABSTRACT: We examine whether short sellers in the equity market provide valuable information to investors in the bond market. Using a sample of publicly traded bond data covering the period from 1988 to 2011, we find that firms with high short interest have lower credit ratings and are more likely to have their ratings downgraded. We also find that firms with highly shorted stocks are associated with higher bond yield spreads (about 24 basis points). Evidence of causality from short interest spikes and a natural experiment based on the SEC's Regulation SHO pilot program confirms our findings. Overall, our results suggest that equity short sellers provide predictive information to creditors in the bond market.

Keywords: *short interest*; *credit ratings*; *bond yield spreads*; *financial reporting.*

Data Availability: *Data are publicly available from the sources identified in the study with the exception of the bond data from Lehman Brothers, which is a proprietary dataset.*

JEL Classifications: *G12; G14.*

I. INTRODUCTION

Equity short sellers play an important role in the informational efficiency of financial markets. They base their investment decisions, in part, on corporate filings that utilize accounting statements, footnotes, and other disclosures.[1] Theory argues that short sellers

The authors thank Bryan Cloyd, Bowie Hanson, Jack Maher, two anonymous referees, John Harry Evans III (senior editor), and especially Steve Kachelmeier (former senior editor) for valuable comments and suggestions that greatly improved the study. We are indebted to Bill Maxwell for providing publicly traded debt data. The remaining errors are the sole responsibility of the authors.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: March 2011
Accepted: September 2012
Published Online: September 2012

[1] Direct evidence comes from short sellers (e.g., see Chanos 2003; Einhorn 2008), from short seller memoirs (e.g., see Schilit 2002), and from the press (e.g., see Staley 1997). According to this evidence, professional short sellers conduct rigorous financial analysis to identify companies that have materially overstated earnings, shifted current expenses to later periods, failed to record or improperly reduced liabilities, and engaged in outright fraud.

help impound negative news more quickly into stock prices (e.g., Diamond and Verrecchia 1987). Consistent with this premise, empirical evidence suggests that equity short sellers are able to profit from their information advantage by selling before decreases in stock prices (e.g., Boehmer et al. 2008). Indeed, it appears that short sellers are able to predict a variety of bad corporate events from changes in analysts' earnings estimates and recommendations to negative earnings surprises, restatements, financial fraud, and accounting irregularities that attract class action litigations.[2] This suggests the disclosure of short interest in stocks can provide investors with information about firms' financial health and viability as well as future securities prices.

In this study, we examine whether short sellers in the stock market provide information to investors in the bond market. We focus on the bond market because of its sheer size and because of its importance in the dissemination of information to the real economy. Creditors in the public bond market rely on accounting-based numbers to monitor their investment, to determine compliance with debt covenants, and to administer lending agreements (DeFond and Jiambalvo 1994; Anderson et al. 2004). This suggests that creditors are concerned over issues that influence the reliability and validity of the financial accounting process. We posit that the information content in short interest (e.g., earnings restatements, earnings quality, auditor choice, and accrual quality) is of great value to creditors and expect debt holders to be sensitive to the firm's level of, as well as changes in, short interest.[3]

We test the hypothesis that an increase in equity short selling increases the cost of debt financing using panel data on a large sample of publicly traded firms covering the period from 1988 to 2011. We begin our analysis by examining the relation between equity short selling and credit ratings as well as bond yield spreads, a more direct measure of the cost of debt financing. We find that firms with highly shorted stocks have significantly lower credit ratings and are more likely to have their ratings downgraded. A one standard deviation increase in short interest increases the probability of a downgrade from 0.48 percent to 0.58 percent in the next month. We also find that firms with high short interest are associated with significantly higher bond spreads. An increase of one standard deviation in short interest is associated with an increase of about 26 basis points in the yield spread. The results are robust to the inclusion of other information intermediaries such as analyst activity and indicate that even when analyst coverage is extensive, short sellers can still fill a complementary information intermediary role (Drake et al. 2011).

Next, we examine whether the relation between short sales and yield spread is causal. To establish causality, we conduct two tests. First, we follow Pownall and Simko (2005) and employ a short-window event study design around short interest spikes. We posit that the announcement of short interest is also informative to bond investors. Bond investors will view a short interest spike as providing new information about the financial health of the firm and thus the default risk of the bond. We identify large and sudden increases and decreases in short interest (i.e., positive and negative short interest spikes) and compute abnormal bond returns around the public disclosure of these spikes. Based on a variety of windows ranging from -1 to $+5$ days, similar to those of the Pownall and Simko (2005) study, we document negative and significant abnormal bond returns

[2] See Francis et al. (2006) and Christophe et al. (2010) for changes in analysts' earnings estimates and recommendations, Pownall and Simko (2005) and Drake et al. (2011) for the interplay between short sellers and research analysts, Christophe et al. (2004) for negative earnings surprises, Desai et al. (2006) for earnings restatements, Karpoff and Lou (2010) for financial fraud, Efendi and Swanson (2009) for accounting irregularities, and Hirshleifer et al. (2011) for operating accruals.

[3] For evidence on the relation between earnings restatements, earnings quality, accrual quality, auditor choice, and the cost of debt, see Shi and Zhang (2008), Qi et al. (2010), Mansi et al. (2004), and Mansi et al. (2011). For evidence on corporate misreporting and bank loan contracting, see Graham et al. (2008); and for evidence on earnings quality and the cost of capital, see Francis et al. (2005b).

(between 2 to 27 bps) for large increases in short interest, and positive and significant abnormal bond returns (between 23 to 47 bps) for large decreases in short interest.

Second, we conduct a natural experiment using the SEC's Regulation SHO pilot program. The pilot program temporarily suspended a short sales constraint for a sample of publicly traded firms. This short sales constraint is binding (e.g., see Alexander and Peterson 1999), and the firms for which it was suspended were chosen randomly by the SEC. Therefore, it is a natural experiment that we can use to study the effect of exogenous changes in equity short selling on the cost of debt. We compare changes in the cost of debt for treatment and control firms and find that the increase in the cost of debt is larger for treatment than for control firms.

We also examine the conditional effect of short sales on the cost of debt financing for smaller firms and those that are financially constrained (i.e., those that do not generate sufficient internal financing and have limited access to the capital market). We posit that bond investors in financially constrained firms are more sensitive to the negative information content of short interest (e.g., poor earnings quality, earnings restatements, and accounting irregularities), which is directly related to the default risk of the bond and the degree of financial constraint.[4] To the extent that equity short selling activity provides a signal about the credibility and quality of financial reporting, we expect the effect of equity short interest on the cost of debt to be higher for financially constrained firms. We find evidence consistent with this reasoning. Again, among small and financially constrained firms, the increase in cost of debt financing for treatment firms is significantly higher than that for control firms.

Our study contributes to the accounting and finance literatures on the information content of short sales, the flow of information between the stock market and the bond market, and alternative sources of information in the bond market. First, we provide new evidence that greater equity short selling causes a higher cost of debt, and thus extend the current literature on the important role short sellers play in the informational efficiency of financial markets. This issue is of importance to financial reporting because it influences whether public disclosure is expected to have significant effects on security prices (Beaver 1998). We argue that short sellers are skilled information processors who not only have the ability to analyze publicly available information to identify overvalued stocks, but can also detect negative events such as accounting irregularities, deteriorating business margins, and corruption among management.

Second, we provide complementary associative evidence that there is information spillover from the stock to the bond market. Existing studies examine the timing of information efficiency in the stock and bond markets, with the majority of research documenting evidence that stock prices lead bond prices.[5] For example, Kwan (1996) finds that stocks incorporate firm-specific information into prices faster than bonds. Gebhardt et al. (2005) find a momentum spillover effect from past equity returns to future bond returns, indicating that bond prices underreact to the information in past stock returns, but not to the information in past bond returns. Downing et al. (2009) find that stock returns predict returns on BB- and non-investment rated bonds at daily and hourly frequencies, suggesting that information in the stock market is valuable to traders in the bond market.[6] We show that it is not just stock prices—set as the equilibrium of both long and short positions—but also the information in short sales that is relevant to investors in the bond market.

[4] Chen et al. (2012) find that firms that restated their earnings are significantly more likely to be financially constrained and have difficulty obtaining new external capital.

[5] One exception is a study by Hotchkiss and Ronen (2002), who find that stock and bond markets are equally informationally efficient.

[6] See also Fleming et al. (1998) for the information flow in the stock, bond, and money markets, and Norden and Weber (2004) for the sensitivity of CDS market to information in the stock and bond markets.

Third, it is generally considered to be the case that the main information producers in the bond market are the credit rating agencies (e.g., Jorion et al. 2005) and the institutional investors themselves that dominate the bond market (e.g., Bessembinder et al. 2005; Edwards et al. 2007). Our study suggests that information from short sellers in the stock market is a significant determinant of bond prices above and beyond information from credit rating agencies and institutional investors. Short sellers, as sophisticated users of financial statements and disclosures, play a valuable role in information revelation that helps provide efficiency to the pricing of securities.

Finally, our study is related to recent research by Henry et al. (2010), who document that short interest increases significantly prior to credit rating downgrades. However, we differ from theirs in that we investigate the relation between relative short interest and the cost of debt financing. We are the first to offer direct evidence that equity short selling is a valuable information source for corporate bond pricings.

Section II provides the data, sample, and summary statistics. Sections III and IV provide the results on the relation between short interest and credit ratings as well as bond yield spreads. Section V establishes causality between short interest and the cost of debt using the Pownall and Simko (2005) event study methodology, and a natural experiment using the SEC's Regulation SHO program. Section VI concludes.

II. DATA AND VARIABLE MEASUREMENT

Data Description

We utilize three main databases in our analysis of the relation between equity short selling and bond pricing: the Lehman Brothers Fixed Income (LBFI) database, the Trade Reporting and Compliance Engine (TRACE) database, and the equity short interest database. The LBFI database contains month-end bond-specific information such as bid price, coupon, yield to maturity, credit ratings from Standard & Poor's (S&P), callability, duration, and maturity dates on nonconvertible bonds that are used in the Lehman Brothers bond indexes. This database covers the period from January 1988 through December 2006 and has been commonly used in the fixed income and corporate finance literatures (e.g., see Anderson et al. 2003; Mansi et al. 2011). The TRACE database consolidates daily transaction data on OTC activity that represents over 99 percent of total U.S. corporate bond market activity in over 30,000 securities. This database covers the period from January 2007 through June 2011.[7] We obtain monthly short interest data from the NYSE, AMEX, and NASDAQ prior to year 2003 and from Compustat thereafter. A firm's short interest is a snapshot of the total number of outstanding short positions on the final trading day on or before the 15th of each month.[8]

We also utilize six supporting databases primarily for controls. These include the Compustat Industrial annual database for financial information, the Center for Research in Security Prices (CRSP) database for stock prices and returns, the Thomson 13F database for institutional holdings, the ExecuComp database for managerial ownership, the RiskMetrics database for board structure variables, and the I/B/E/S database for stock analyst characteristics.

---

[7] Although the TRACE database provides data from July 2002 to June 2011, for the purpose of our main analysis we only use data from TRACE when the LBFI dataset is no longer available (i.e., December 2006).

[8] Our sample period includes the recent financial crisis when the SEC imposed several restrictions on equity short selling. To ensure that our results are not driven by this financial shock, we use data before the financial crisis and find similar results.

For an observation to be included in the analysis, firms must have a debt issue that is present in the LBFI and TRACE datasets, and have available equity short interest information and firm-specific characteristics. We include only common stocks (share code 10 or 11) and exclude financial firms with SIC codes between 6000 and 6999 because they differ in their treatment of certain firm-specific variables such as financial leverage. To minimize pricing errors, we also exclude all matrix-priced bonds, AAA-rated bonds, zero-coupon bonds, floating rate debt, bonds with zero or negative yield spread, bonds with odd frequency coupon payments, and bonds with less than one year to maturity due to their illiquidity (e.g., see Campbell and Taksler 2003; Bharath and Shumway 2008). Merging the data and applying these requirements yields a dataset of 10,594 firm-year observations on 1,714 firms covering a sample period from January 1988 through June 2011.

Measuring Short Interest and Bond Yield Spread

The primary variable used in this study is relative short interest (*RSI*), defined as the number of shares shorted scaled by the number of shares outstanding. This variable has been widely used in prior studies to proxy for monthly equity short selling activities (e.g., Drake et al. 2011; Pownall and Simko 2005; Dechow et al. 2001).

To compute bond yield spread for each corporate debt instrument, we use the yield to maturity from the LBFI and the TRACE databases. The yield spread is measured as the difference between the yield to maturity on a corporate bond, or the discount rate that equates the present value of its future cash flows to its current price, and the yield to maturity on its duration-equivalent Treasury security. The yield to maturity on a Treasury security is the yield on the constant maturity series obtained from the Federal Reserve Bank in its H15 release based on a par bond. In the cases where no corresponding Treasury yield is available for a given maturity, we calculate the yield spread using linear interpolation. For firms with multiple bonds, we compute a weighted average yield spread, with the weight being the offering amount of the bond divided by the total number of offering amounts for all available bonds.[9]

The central objective of this study is to examine whether short sellers provide valuable information for bond pricings. Therefore, in each year we use the month of June as a snapshot to structure our dataset. In particular, we only keep bonds in the June month of each year *t* so that each bond will only be present once in our sample every year. We use short interest information in the prior month (May) of year *t* to construct our *RSI* variable, stock information up to the month of May of year *t* to compute stock-specific variables, and accounting information in the fiscal year that ends in *t*–1 to compute accounting variables. This ensures that short selling information in the equity market and firm fundamentals are both available to the bond investors.

Control Variables

The remaining variables are firm- and security-specific controls. Firm-specific controls include firm size, market-to-book, momentum, leverage, profitability, tangibility, stock volatility, capital expenditures, and firm age. Firm size (*Size*), a proxy for economies of scale and a takeover deterrent, is the natural log of book value of total assets. The market-to-book ratio (*Market-to-Book*), a proxy for growth opportunities, is the market value of equity scaled by book equity. Stock momentum (*Momentum*), a proxy for spillover effect, is the 12-month buy-and-hold stock return. Firm leverage (*Leverage*), a proxy for financial health, is the ratio of long-term debt scaled by book value of assets. Firm profitability (*Profitability*), a proxy for financial performance, is the ratio of

[9] We also use each bond as an independent observation and find similar results.

TABLE 1

Variable Definitions

Main Variables	Description	Database
Relative Short Interest	Number of shares short sold scaled by the number of shares outstanding	[a]
Yield Spread	The difference between the yield to maturity on a corporate bond and the yield to maturity on its duration equivalent Treasury security	LBFI, TRACE
Market Capitalization	Equity market capitalization (in $millions)	Compustat
Firm Size	Log of book value of assets (in $millions)	Compustat
Market-to-Book	Log (Number of shares outstanding times the stock price scaled by book value of equity)	Compustat
Momentum	12-month buy-and-hold stock returns	CRSP
Profitability	Earnings before interest, tax, depreciation, and amortization scaled by firm sales	Compustat
Tangibility	Property, plant, and equipment scaled by book value of assets	Compustat
Volatility	Annualized standard deviation of 12-month daily returns	CRSP
Capx	Capital expenditure scaled by book value of asset	Compustat
Firm Age	Log number of years since firm appears for first time in the CRSP tape	CRSP
Inst-Own	Number of shares held by institutions scaled by common shares outstanding	Thomson 13F
Mgr-Own	Number of shares held by executives scaled by common shares outstanding	ExecuComp
Board Size	Log of number of board members	RiskMetrics
Independence	Number of outside board directors scaled by number of total board members	RiskMetrics
Analyst	Log of number of analysts following the stock	I/B/E/S
Dispersion	Standard deviation of analyst forecast scaled by mean absolute forecast	I/B/E/S
Rating	S&P bond rating	LBFI
Bond Return	Change in price plus accrued interest scaled by initial period price	TRACE
Bond Age	Log of number of years since bond issuance	LBFI, TRACE
Duration	Macaulay duration or security's effective maturity	LBFI, TRACE
Leverage	Total Debt scaled by book asset	Compustat
High Yield	A dummy variable that equals 1 if the debt is non-investment grade	FISD

[a] Data before 2003 are from exchanges, and data from 2003 and onward are from Compustat.
LBFI is the Lehman Brothers Fixed Income database. TRACE is the Trade Reporting and Compliance Engine database provided by the NASD.

earnings before interest, tax, depreciation, and amortization scaled by firm sales. Firm tangibility (*Tangibility*) is the ratio of property, plant, and equipment scaled by book value of assets. Firm volatility (*Volatility*) is the annualized standard deviation of 12-month daily returns, as in Campbell and Taksler (2003). Capital expenditures (*Capx*), a proxy for firm expansion or replacement of physical assets, is the ratio of capital expenditure scaled by book value of asset. Firm age (*Firm Age*), a proxy for stability, is the log of number of years since the firm first appeared in the CRSP tape. Table 1 provides definitions for the variables used in the analysis.

Debt-specific characteristics include credit rating from S&P (*Rating*), debt duration (*Duration*), and debt age (*Bond Age*). The S&P credit rating is computed using a conversion process in which AAA-rated bonds are assigned a value of 22 and non-rated bonds receive a value of zero (see e.g., Anderson et al. [2003] for S&P bond rating conversion numbers). An alternative methodology used in the literature allows us to assume that the credit rating variable may incorporate part or all of the short seller's information. As such, we estimate the effect of credit rating on yield spreads without the relative short interest variable. That is, we regress credit ratings on the relative short interest variable, and the error term in this case incorporates the credit rating information without the influence or impact of short selling. The error term from this regression is labeled as "Credit Rating" and serves as our primary measure of debt rating in the multivariate analysis on yield spreads.

We also allow for a nonlinear relation between bond yield spreads and credit ratings. When examining the entire LBFI dataset, we find that as firm credit ratings move from investment (debt with rating greater than 12) to non-investment-grade debt (debt ratings less than 13), the increase in yield spread for the non-investment categories becomes nonlinear. Therefore, we use an indicator variable (*High Yield*) that equals 1 when the debt is non-investment grade to capture any nonlinearities in the yield spread.

Duration, or effective maturity, is the present value of the weighted average of cash flows divided by the security's price, which we include to control for term structure effects. For liquidity, we use the log of debt age (*Age*), where debt age is the number of years that a bond has been outstanding. Beim (1992) finds that liquidity is positively priced in the debt market as more recently issued bonds are more liquid than older bonds. For firms with multiple bonds, we compute weighted average durations and bond age using the summation of the weighted durations and debt ages of all bonds for each firm, with the weight being the offering amount for each debt issue divided by total offering amount for all publicly traded debt.

We also control for various governance structures that are likely to reduce agency problems and have an effect on bondholder wealth. These include institutional ownership, managerial holdings, and board structure characteristics. Institutional ownership (*Inst-Own*), a proxy for monitoring, is the ratio of shares that institutions owned for a firm divided by the number of shares outstanding (Bhojraj and Sengupta 2003).[10] Managerial ownership (*Mgr-Own*), a proxy for ownership concentration and risk, is measured as the number of shares held by top executives scaled by common shares outstanding. Because insider ownership is often shown to have a nonlinear impact (McConnell and Servaes 1990), we control for managerial ownership and the square of that term. Board characteristics that have been shown to impact yield spreads due to monitoring (Anderson et al. 2004) include board size (*Board Size*), computed as the log of number of board members, and board independence (*Independence*), measured as the ratio of the number of outside board directors scaled by the number of total board members.

We further control for the information environment of the firm. Mansi et al. (2011) find that the log of number of analysts following the stock (*Analyst*) and the standard deviation of analyst forecasts scaled by mean absolute forecast (*Dispersion*) provide information to the debt market beyond that incorporated in firm fundamentals. Drake et al. (2011) argue that stock analysts and short sellers provide information that predicts future returns. As such, controlling for these variables provides us with a further test that equity short selling contains information to bond investors beyond that provided by these intermediaries.

---

[10] This variable is also a proxy for the supply of shares available for shorting (Asquith et al. 2005) and thus may affect the relation between equity short interest and corporate bond yield spreads.

Descriptive Statistics

Table 2, Panel A reports summary statistics for all variables, which are Winsorized at 1 and 99 percentiles to reduce the effect of outliers. The mean, median, and standard deviation of the yield spread are 280, 190, and 283 basis points, with upper and lower quartile values of 340 and 119 basis points, respectively. On average, firms in the sample have relative short interest ratios of 3.0 percent, median ratios of 1.9 percent, with standard deviations of 3.3 percent.

For the overall sample, market capitalization has a mean of \$6.8 billion, a standard deviation of \$12.5 billion, and 75th and 25th percentile values of \$6.6 billion and \$808 million, respectively. The mean and median leverage ratio is close to 34 percent with a standard deviation of 15 percent, which suggests that a large portion of the sample consists of firms with significant long-term debt in their capital structure. The firms in the sample have a median age of 27.4 years, tangible assets of 36 percent, profitability ratio of 4.4 percent, an annualized standard deviation of daily stock returns of 34.2 percent, momentum of 9 percent, and capital expenditures of 5.6 percent. Institutions, on average, own 60.9 percent of the firms' shares outstanding with a standard deviation of 19 percent, while top managers own 2.4 percent of stock in their firm with a standard deviation of 5.5 percent. The average firm has about 11 board members, and two out of three members are independent. The average firm also has about 11 analysts following their stock with an earnings forecast dispersion of 15.4 percent.

In terms of debt variables, the mean (median) bond rating variable equates to S&P ratings of BBB– (BBB) with a standard deviation between A and B+, indicating that a large portion of the sample is on the border between investment and non-investment-grade debt. The mean traded debt has a duration of 5.9 years with a standard deviation of 2.2 years, and on average has been outstanding for more than five years.

Table 2, Panel B describes the industry distribution of the sample (in absolute number and in percentage) using the Standard Industrial Classification (SIC) codes. Industries include: agriculture, forestry, and fishing; mining and construction; manufacturing (food through petroleum, and plastics through electronics); transportation and communications; wholesale and retail trade; services (hotels through recreation, and health through private households); and public administration. Based on the sample, a large portion is concentrated in manufacturing (49 percent), transportation and communication (19 percent), wholesale and retail trade (12 percent), mining and construction (10 percent), and services (11 percent).

Table 3 provides selected correlation coefficients among the relative short interest variable, yield spread, firm-specific variables, and debt characteristics. The yield spread is positively correlated with short interest, leverage, stock volatility, and forecast dispersion, but negatively correlated with firm size, momentum, profitability, credit ratings, institutional ownership, and analyst coverage. In general, the analysis indicates that firms with higher short selling have higher cost of debt financing, evidence consistent with the hypothesis that equity short sellers provide valuable information to the bond market. However, because many firm- and issue-specific characteristics also have an effect on short interest and bond yield spreads, we use a multivariate framework to explore our hypotheses in the following sections.

III. RESULTS ON CREDIT RATINGS

To examine whether short sellers provide valuable information to the corporate bond market, we commence our analysis by focusing on the relation between relative short interest and credit ratings. That is, we examine the relation between relative short interest and a firm's bond rating while controlling for firm-specific factors using two-way clustered standard errors at the year and

TABLE 2
Summary Statistics

Panel A: Descriptive Statistics

Variable	Mean	Median	Standard Deviation	First Quartile	Third Quartile
Relative Short Interest (RSI)	0.030	0.019	0.033	0.009	0.038
Yield Spread (%)	2.804	1.896	2.828	1.188	3.397
Market Capitalization ($B)	6.794	2.306	12.512	0.808	6.602
Total Asset ($B)	8.804	3.424	15.296	1.457	9.116
Market-to-Book	3.109	1.757	10.202	1.160	2.826
Momentum	0.118	0.090	0.341	−0.085	0.278
Leverage	0.346	0.332	0.150	0.240	0.435
Profitability	0.027	0.044	0.153	0.012	0.080
Tangibility	0.407	0.360	0.243	0.205	0.608
Volatility	0.380	0.342	0.164	0.271	0.445
Capx	0.081	0.056	0.087	0.034	0.091
Firm Age	31.279	27.375	21.613	13.095	45.891
Board Size	10.509	10.364	2.320	8.909	12.000
Independence	0.675	0.706	0.172	0.554	0.816
Inst-Own	0.609	0.636	0.189	0.495	0.745
Mgr-Own	0.024	0.004	0.055	0.002	0.013
Analyst	10.675	10.000	7.007	5.042	15.500
Dispersion	0.154	0.052	0.361	0.026	0.126
Bond Age	5.138	3.763	4.330	2.042	7.101
Duration	5.898	5.619	2.192	4.392	7.184
Rating	BBB−	BBB	A/B+	BB−	A−

Panel B: Industry Classifications

SIC Code	Industry Classification	Observations	Percentage (%)	Cumulative (%)
0	Agriculture, Forestry, and Fishing	10	0.09	0.09
1	Mining and Construction	1,046	9.87	9.97
2	Manufacturing (Food-Petroleum)	2,445	23.08	33.05
3	Manufacturing (Plastics-Electronics)	2,699	25.48	58.52
4	Transportation and Communication	2,037	19.23	77.75
5	Wholesale Trade and Retail Trade	1,231	11.62	89.37
7	Services (Hotels-Recreation)	754	7.12	96.49
8	Services (Health-Private Household)	359	3.39	99.88
9	Public Administration and Other	13	0.12	100.00

Industry classifications for the sample are based on one-digit SIC codes. The data comprise 10,594 bonds on 1,714 non-financial firms for the period from January 1988 to June 2011.
Variable definitions are provided in Table 1.

TABLE 3

Correlations among Cost of Debt, Firm Variables, and Credit Rating
(n = 10,594)

	RSI	*Yield Spread*	*Firm Size*	*Leverage*	*Profitability*	*Volatility*	*Momentum*	*Inst-Own*	*Analyst*	*Dispersion*
Yield Spread	0.221**									
Firm Size	−0.127	−0.210**								
Leverage	0.138	0.371***	−0.075							
Profitability	−0.123	−0.353***	0.044	−0.192*						
Volatility	0.245**	0.655***	−0.202*	0.283**	−0.399***					
Momentum	−0.032	−0.182*	−0.011	−0.059	0.101	−0.102				
Inst-Own	0.183*	−0.187*	−0.033	−0.217*	0.049	−0.105	0.099			
Analyst	−0.062	−0.349**	0.325**	−0.274*	0.137	−0.252*	0.189	0.184		
Dispersion	0.143	0.334**	−0.031	0.167**	−0.232**	0.308***	−0.166	−0.035	−0.087	
Rating	−0.283***	−0.648***	0.334***	−0.458***	0.348***	−0.607***	0.047	0.091*	0.427	−0.287***

*, **, *** Significance at the 10 percent, 5 percent, and 1 percent levels, respectively.
Numbers reported are Pearson correlation coefficients. The data cover the period from January 1988 to June 2011.
Variable definitions are provided in Table 1.

firm levels as in Gow et al. (2010). This method controls for potential cross-sectional correlation and intertemporal dependence among the residuals. Our primary specification is:

$$Rating_{i,t} = \beta_0 + \beta_1(RSI_{i,t-1}) + \beta_{2-9}(Firm\ Specific\ Factors_{i,t-1}) + \beta_{10}(Inst\text{-}Own_{i,t-1}) + \beta_{11-16}(Industry\_Dum_{i,t-1}) + \varepsilon_{i,t}, \quad (1)$$

where *Rating* is the credit rating variable, *RSI* is relative short interest, *Firm Specific Factors* include firm size, firm age, leverage, profitability, market-to-book, momentum, tangibility, volatility, institutional ownership, and capital expenditures. *Industry_Dum* represents industry indicator variables. Our principal concern in the analysis is the relative short interest coefficient estimate, β_1. Because both the governance structure and the information environment of the firm are relevant to the pricing of debt securities, we include managerial ownership, analyst coverage, forecast dispersion, board size, and board independence as additional control variables in our regressions. We expect the variables that reduce agency costs (e.g., managerial holdings, board structure) and those that provide additional information to the market (e.g., analyst forecast characteristics) to be associated with better credit rating.

Table 4, Panel A provides results of our credit rating analysis. Model 1 is our primary specification. Models 2, 3, and 4 are similar to the primary specification but with additional controls for managerial ownership, information environment (financial analyst following and forecast dispersion), and board structure variables (board size and board independence), respectively. Model 5 provides the full specification with all of the control variables from prior models included.

Across all models, we find a statistically significant (at $p < 0.01$ level) inverse relation between prior short interest and firm credit ratings after controlling for firm- and debt-specific characteristics, governance, and information-related variables. The coefficients across models vary and suggest that firms with higher short interest have ratings that are 0.57 to 0.74 of a rating step lower than firms with low short interest. These models explain over 53 percent of the cross-sectional variation of firms' bond ratings and suggest that short sellers provide information to the credit rating market beyond that incorporated in firm governance and analyst forecasts.[11]

We also examine the informativeness of annual change in *RSI* in predicting the annual change in credit rating in Panel B of Table 4 using the same specifications as in Panel A. The dependent variable is the change in bond rating from June of year *t*–1 to June of year *t*, and the variable of interest is the change in *RSI* from May of year *t*–1 to May of year *t*, with the annual changes in all controls also included in the regression. The coefficients on the change in *RSI* across all models vary from –3.1 to –5.3, results that are statistically significant at $p < 0.01$. The control variables have their expected signs and in general are statistically significant. Specifically, we find that firms with better credit ratings are more profitable and have less leverage, greater tangible assets, lower momentum, and less earnings volatility. These firms are also large and have been in the market for a long period of time. Overall, the results indicate that short selling provides important information about future levels as well as changes in credit ratings.

We further examine whether short interest is informative about credit rating downgrades over a shorter horizon window. We estimate a binominal logit regression to determine the relative importance of *RSI* in predicting downgrades in the immediate next month. The idea is that sophisticated equity investors such as short sellers may have negative information about the firm and thus anticipate bond downgrades. The dependent variable is an indicator variable that equals 1 if the firm bond rating is downgraded in month *t*. The independent variables *RSI*, size, leverage,

[11] Note that the adjusted R^2 decreases from the base model to the other four models in Table 4, Panel A, even though more variables are added to the model due to the reduction in sample size across the models.

TABLE 4

Relative Short Interest and Firm Credit Rating

Panel A: Firm-Level Specifications

	Primary Specification (1)	Managerial Ownership (2)	Information Variables (3)	Board Structure (4)	Full Specification (5)
RSI	−24.693***	−21.923***	−21.151***	−20.628***	−18.873***
	(−12.00)	(−11.16)	(−8.60)	(−7.19)	(−6.26)
Firm Size	1.105***	0.985***	0.811***	0.748***	0.614***
	(11.56)	(8.74)	(6.17)	(8.91)	(5.20)
Firm Age	0.609***	0.487***	0.600***	0.321***	0.336***
	(8.85)	(5.86)	(9.08)	(3.45)	(3.56)
Leverage	−5.911***	−5.463***	−6.070***	−6.122***	−5.893***
	(−8.37)	(−7.60)	(−9.87)	(−10.40)	(−9.76)
Profitability	2.733***	2.654***	2.398***	3.150***	2.424**
	(8.05)	(4.55)	(6.38)	(5.74)	(2.21)
Market-to-Book	0.745***	0.907***	0.499***	1.040***	0.848***
	(7.04)	(8.17)	(4.97)	(10.21)	(7.13)
Momentum	−0.676***	−0.728***	−0.600***	−0.772***	−0.765***
	(−3.50)	(−3.41)	(−3.91)	(−4.52)	(−3.78)
Tangibility	1.384***	0.982**	0.749*	0.166	−0.098
	(4.07)	(2.39)	(1.89)	(0.37)	(−0.21)
Volatility	−4.610***	−4.427***	−3.695***	−3.235***	−3.224***
	(−8.86)	(−8.89)	(−6.83)	(−5.15)	(−5.61)
Inst-Own	−2.452***	−3.308***	−2.409***	−1.812**	−2.621***
	(−5.90)	(−6.87)	(−5.30)	(−2.41)	(−3.83)
Capx	1.434*	1.668	0.911	2.095	2.037
	(1.75)	(1.45)	(1.28)	(1.53)	(1.46)
Mgr-Own		−1.278			−3.915
		(−0.35)			(−0.90)
Mgr-Squared		2.219			8.855
		(0.18)			(0.55)
Analyst			1.025***		0.541***
			(7.11)		(3.95)
Dispersion			−1.089***		−0.898***
			(−8.00)		(−4.06)
Board Size				1.908***	1.462***
				(4.58)	(3.38)
Independence				0.187	−0.060
				(0.42)	(−0.14)
Adj. R^2	0.590	0.531	0.626	0.540	0.541
Observations	10,594	7,469	7,685	3,726	3,050

(continued on next page)

TABLE 4 (continued)

Panel B: First-Difference Regressions

	Primary Specification (1)	Managerial Ownership (2)	Information Variables (3)	Board Structure (4)	Full Specification (5)
RSI	−3.065***	−4.245***	−5.282***	−4.839***	−4.897***
	(−3.64)	(−4.59)	(−5.46)	(−3.22)	(−2.68)
Firm Size	0.685***	0.802***	0.602***	0.884***	0.803***
	(6.70)	(9.95)	(4.89)	(3.43)	(2.87)
Firm Age	0.187	0.284**	0.341	−0.193	−0.446
	(1.07)	(2.25)	(1.47)	(−0.48)	(−0.84)
Leverage	−1.767***	−2.185***	−1.741***	−2.035***	−2.423***
	(−5.92)	(−5.69)	(−3.82)	(−4.13)	(−3.70)
Profitability	0.324	0.257	0.803***	1.163**	1.569***
	(1.67)	(1.26)	(3.37)	(2.24)	(2.66)
Market-to-Book	0.185***	0.240***	0.137***	0.303***	0.246**
	(4.60)	(4.67)	(2.74)	(4.06)	(2.43)
Momentum	−0.133***	−0.142**	−0.168***	−0.113	−0.095
	(−2.91)	(−2.39)	(−3.58)	(−1.26)	(−1.31)
Tangibility	0.891*	1.301**	1.121	1.111	1.633
	(1.65)	(2.07)	(1.62)	(0.73)	(1.00)
Volatility	−0.950***	−0.949***	−0.913**	−1.422*	−1.226*
	(−4.18)	(−3.54)	(−2.38)	(−1.85)	(−1.67)
Inst-Own	0.190	0.591*	0.228	0.446	0.561
	(0.77)	(1.76)	(0.73)	(0.91)	(0.79)
Capx	0.280	0.131	0.303	0.982	0.817
	(0.90)	(0.25)	(0.92)	(1.61)	(1.44)
Mgr-Own		2.307			5.456*
		(1.34)			(1.77)
Mgr-Squared		−6.421			−15.293*
		(−1.27)			(−1.85)
Analyst			0.008		0.168*
			(0.11)		(1.71)
Dispersion			−0.164***		−0.174***
			(−2.56)		(−2.60)
Board Size				0.309	0.259*
				(1.30)	(1.65)
Independence				0.026	0.094
				(0.08)	(0.36)
Adj. R^2	0.054	0.064	0.062	0.070	0.082
Observations	8,177	5,846	5,559	2,780	2,191

*, **, *** Significance at the 10 percent, 5 percent, and 1 percent levels, respectively.
Panel A regresses credit rating in June of year t on *RSI* and stock-specific variables in May of year t and accounting variables in year t−1. Panel B regresses change in credit rating from June of year t−1 to June of year t on change on *RSI* and stock-specific variables from May of year t−1 to May of year t and change in accounting variables in year t−2 and year t−1. The data cover the period from January 1988 to June 2011.
See Table 1 for variable definitions. All models include industry fixed effects. t-statistics (in parentheses) are adjusted using firm and year two-way cluster-robust standard errors as in Gow et al. (2010).

profitability, market-to-book, momentum, volatility, and institutional ownership are all constructed in month t–1.[12]

We consider three different cases of bond downgrades: (1) all downgrades, (2) firms whose bond credit rating is downgraded from investment to non-investment-grade category, and (3) firms whose credit rating is downgraded across major rating categories (i.e., from AA to BBB or below, from BBB to BB or below). In our sample of 130,992 firm-month observations, we have 4,260 cases of downgrades, of which 2,209 are within rating categories, 1,718 across rating categories, and 333 from investment to non-investment-grade debt.

The results are presented in Table 5 and suggest that *RSI* is a significant predictor of future downgrades in all models. If we set all controlling variables at their sample median, a one standard deviation increase in *RSI* will increase the probability of a downgrade in the following month from 0.93 percent to 1.10 percent in Model 1, from 0.14 percent to 0.16 percent in Model 2, and from 0.85 percent to 1.02 percent in Model 3. Overall, these models have good discriminatory power in inferring downgrades from *RSI*. In particular, the area under the receiver operator characteristic (ROC) curve for these models is in the range of 0.65 to 0.68. Thus, there is a 65 percent to 68 percent chance that a firm with high *RSI* has a higher probability of credit rating downgrade in the subsequent month than a firm with low *RSI*. Because large and growth firms that already have good credit quality have limited potential to raise their credit quality but a greater potential for being downgraded, we find a significant and positive coefficient on firm size in all models. Recently, Henry et al. (2010) document that, in the month preceding a credit rating downgrade, equity short interest is 40 percent higher than one year prior. Our finding offers further support for their argument that short sellers provide a role in identifying firms that are likely to be downgraded. Overall, short sellers in the stock market provide valuable information about bond future credit rating downgrades.

IV. RESULTS ON BOND YIELD SPREADS

We now investigate the relation between relative short interest and bond yield spreads. Because rating agencies and short sellers are information providers, the direction of causality between them is not obvious and may be determined endogenously. In addition, since credit ratings do not change frequently, first-difference models may not be appropriate. In fact, even if we assume that equity short sellers can "cause" credit ratings to change, it is not necessary that this information trickles down to bond yields in the same way that the credit rating agencies use them when they set the ratings. Prior research also provides evidence that the market impounds information to stock prices well before the information is reflected in bond ratings (e.g., see Dichev and Piotroski 2001). Therefore, examining yield spreads allows us to directly study how the market values the information that short sellers provide while controlling for any additional information that the rating agencies provide.

We first examine the association between relative short interest and bond yield spreads while controlling for factors that are known to influence yield spreads. We perform multivariate analysis using pooled cross-section and time-series regressions, firm fixed effects, and first difference specifications. With the exception of the firm fixed effect model, we use two-way clustered standard errors at the year and firm levels as in Gow et al. (2010). Our primary regression model is:

$$\begin{aligned} Spread_{i,t} = \beta_0 + \beta_1(RSI_{i,t-1}) + \beta_{2-9}(Firm\ Specific_{i,t-1}) + \beta_{10}(Inst\text{-}Own_{i,t-1}) \\ + \beta_{11-14}(Security\ Specific_{i,t-1}) + \beta_{15-20}(Industry\_Dum_{i,t-1}) + \varepsilon_{i,t}, \end{aligned} \tag{2}$$

[12] In unreported regressions, we use short interest in June of year t to forecast various downgrades during the next 12 months and find similar results.

TABLE 5

Short Interest and Probability of Credit Downgrades

	All Downgrades (1)	Non-Investment Grade Downgrades (2)	Category Downgrades (3)
RSI	5.154***	3.210***	5.421***
	(< 0.01)	(0.01)	(< 0.01)
Firm Size	0.249***	0.287***	0.083***
	(< 0.01)	(< 0.01)	(< 0.01)
Leverage	0.711***	−1.209***	0.549***
	(< 0.01)	(< 0.01)	(< 0.01)
Profitability	0.098	0.073	0.184
	(0.25)	(0.85)	(0.12)
Market-to-Book	−0.159***	−0.283***	−0.125***
	(< 0.01)	(< 0.01)	(< 0.01)
Momentum	−0.658***	−0.892***	−0.672***
	(< 0.01)	(< 0.01)	(< 0.01)
Volatility	1.136***	−0.175	1.224***
	(< 0.01)	(0.57)	(< 0.01)
Inst-Own	0.202**	0.239	−0.121
	(0.01)	(0.39)	(0.30)
Observations	130,992	130,992	130,992
Downgrade Obs.	4,260	333	1,718
Pseudo R^2	0.035	0.002	0.006
Area under ROC Curve	0.681	0.654	0.674

, * Significance at the 5 percent and 1 percent levels, respectively.
Logit regressions of the downgrade indicator are in month $t+1$, and *RSI* and other firm-specific variables all constructed in month t. The data cover the period from January 1988 to June 2011. The downgrade indicator equals 1 if the firm's bond rating is downgraded in month $t+1$.
Variable definitions are included in Table 1. ROC is the receiver operating characteristics curve. All models include industry fixed effects. The p-values are reported in parentheses below each coefficient estimate.

where *Spread* is the yield spread and *RSI* is relative short interest. *Firm Specific* factors include firm size, age, leverage, profitability, market-to-book, momentum, tangibility, volatility, capital expenditures, and institutional ownership. *Security Specific* factors include credit ratings, debt duration, debt age, and high-yield indicator. Our principal concern in the analysis is the relative short interest coefficient estimate, β_1. A significant and positive coefficient would provide support for the hypothesis that equity short selling provides value-relevant information to bond pricings.

For our firm-specific control variables, we expect both firm size and age to be negatively related to yield spread as larger firms enjoy economies of scale and greater stability. We also expect firms with high past stock returns to have lower yield spreads. Leverage should be positively related to yield spreads, as higher debt capacity is associated with higher probability of default. The market-to-book ratio should be negatively associated with yield spreads as firms with higher growth opportunities utilize less debt and therefore lower probability of default. We expect capital investment and firm profitability to be negatively related to the cost of debt financing, as more profitable firms usually invest more (Morck et al. 1990) and have a lower probability of default.

Table 6, Panel A provides our primary specifications. Model 1 reports our main results. Models 2 and 3 utilize first difference and firm fixed effect regressions. Models 4 and 5 control for the

TABLE 6

Relative Short Interest and Yield Spread

	Baseline (1)	First Differences (2)	Fixed Effects (3)	Information Variable (4)	Interaction with Info. Var. (5)
RSI	7.919***	3.629**	5.066***	6.857***	7.659*
	(4.99)	(2.14)	(5.74)	(4.62)	(1.82)
Firm Size	−0.023	0.006	−0.184***	0.016	0.015
	(−0.36)	(0.25)	(−3.01)	(0.26)	(0.25)
Firm Age	−0.032	1.583**	0.540***	−0.008	−0.009
	(−0.71)	(2.29)	(5.35)	(−0.21)	(−0.26)
Leverage	1.103***	2.211***	1.281***	0.775**	0.776**
	(3.10)	(2.94)	(4.69)	(2.00)	(1.99)
Profitability	−1.315***	−0.817	−0.649***	−1.609**	−1.594**
	(−2.62)	(−1.52)	(−3.71)	(−2.29)	(−2.25)
Market-to-Book	−0.213***	−0.502***	−0.357***	−0.163**	−0.165**
	(−2.72)	(−2.75)	(−8.06)	(−2.09)	(−2.06)
Momentum	−0.941***	−0.768***	−0.963***	−0.933***	−0.930***
	(−4.64)	(−6.24)	(−16.7)	(−6.18)	(−6.12)
Tangibility	−0.379*	1.738*	1.085***	−0.072	−0.074
	(−1.83)	(1.91)	(3.39)	(−0.37)	(−0.39)
Volatility	6.392***	3.742***	6.521***	5.592***	5.601***
	(12.19)	(3.41)	(33.34)	(8.07)	(8.14)
Inst-Own	−0.763**	−1.618***	−2.224***	−0.396	−0.395
	(−1.96)	(−2.99)	(−10.27)	(−1.25)	(−1.25)
Capx	−0.730	−1.577	−1.084***	−0.341	−0.333
	(−0.94)	(−1.15)	(−2.86)	(−0.46)	(−0.45)
Rating	−0.110***	−0.139***	−0.102***	−0.099***	−0.099***
	(−5.16)	(−2.94)	(−7.24)	(−4.48)	(−4.54)
High Yield	0.874***	0.452**	0.784***	0.883***	0.887***
	(5.00)	(2.04)	(7.67)	(6.66)	(6.72)
Duration	−0.149***	−0.394***	−0.063***	−0.045	−0.045
	(−2.67)	(−2.96)	(−4.78)	(−1.37)	(−1.38)
Bond Age	0.145***	0.002	0.167***	0.124***	0.123***
	(4.30)	(0.03)	(6.56)	(3.16)	(3.11)
Analyst				−0.293**	−0.276**
				(−2.53)	(−2.08)
Dispersion				0.314*	0.242
				(1.75)	(1.38)
RSI × *Analyst*					−0.602*
					(−1.92)
RSI × *Dispersion*					1.361**
					(2.13)
Adj. R^2	0.528	0.264	0.542	0.534	0.534
Observations	10,594	8,689	10,594	7,685	7,685

*, **, *** Significance at the 10 percent, 5 percent, and 1 percent levels, respectively.
The data cover the period from January 1988 to June 2011.
Variable definitions are provided in Table 1. t-statistics (in parentheses) are adjusted using firm and year two-way cluster-robust standard errors as in Gow et al. (2010) in all models except Model 3, where we conduct a panel regression with fixed firm and year effects.

information environment and the interaction of the information variables with relative short interest. In all models, we control for credit ratings using the residual from the regression of S&P bond rating on relative short interest. Across all specifications, we find a positive and significant relation between relative short interest and bond yield spread, indicating that higher short interest is associated with higher future cost of borrowing. The results in the baseline specification translate to an increase in bond yield spread of about 24 basis points for the average firm. For two otherwise identical firms, an increase of one standard deviation in *RSI* is associated with an economic increase of about 26 basis points in yield spread.[13]

We also examine the relation between short selling and bond yield spread while controlling for other information intermediaries such as equity analysts. Mansi et al. (2011) find that stock analysts provide information that is valued outside the equity market in which better information is associated with a lower cost of debt capital. Drake et al. (2011) find that short sellers have a complementary information intermediary role even when coverage by financial analysts is extensive. Pownall and Simko (2005) find that the information content of short interest disclosures is conditional on firms' information environment. Since short sellers and analysts, as information providers, invest considerable time and resources analyzing companies, an interesting question is whether short sellers provide information above and beyond that supplied by equity analysts.[14]

We use two variables to proxy for the information environment: analyst following and forecast dispersion. More analyst followings indicate more coverage and hence more available information about the firm, while higher forecast dispersion indicates greater information asymmetry and thus lower information quality. The results are provided in Model 4. Similar to our earlier findings, we find a positive and significant (at $p < 0.01$) relation between short selling and bond yield spread (about 21 basis points), suggesting the equity short selling effect on yield spreads is not subsumed by the information environment of the firm.

To examine whether analysts and short sellers provide complementary roles, we interact *RSI* with analyst following and forecast dispersion in Model 5. The coefficient estimates on the two interaction terms represent the conditional effects of *RSI* on bond pricings for a given level of information quantity and quality produced by equity analysts. Compared to Model 4, analyst following continues to have a significant and negative coefficient (at $p < 0.05$), but the coefficient estimate on forecast dispersion becomes insignificant. We find that the coefficient estimates on the two interaction terms are both significant, suggesting that the *RSI* effect on yield spreads depends on the information environment. In particular, while unconditionally yield spreads are positively related to *RSI*, less analyst following and high forecast dispersion further magnify this positive relation. Pownall and Simko (2005) document that short sellers serve as important information intermediaries to stock investors, with the perceived value of short sellers highest when there are limited alternative sources of guidance from equity analysts. We find that the strength of the short selling signal to bond investors is decreasing in the quantity and quality of information provided by equity analysts, evidence that is broadly consistent with Pownall and Simko (2005). The results indicate that short sellers provide valuable information to the bond market, especially when analyst forecasts are not extensive.

To address the concern that a correlated omitted variable may be driving our results, we also examine the relation between short selling and bond yield spreads using a first-difference specification (Model 2). Similar to our level of specification, we find a positive and significant relation (at $p < 0.05$) between changes in short selling and changes in yield spreads. The results are

---

[13] We also examine the relation between *RSI* and the yield spread using the log transformed variable as the dependent variable and find similar results.

[14] Note that short sellers have a different incentive structure than analysts in that they place their own capital at risk and therefore have a strong desire to fully use predictive information.

also economically significant: an increase of one standard deviation in *RSI* from prior year is associated with an annual increase of yield spread of about 12 basis points. Also, given the potential for the existence of a spurious relation between *RSI* and yield spread, it is difficult to adequately control for the simultaneity between them. Nevertheless, estimation employing alternative methodology provides some enhancement to the integrity of the analysis (Coles et al. 2012). Therefore, we use a fixed effects model that relies on time series variations to identify the relation between yield spreads and *RSI* in Model 3. The *RSI* coefficient in the fixed effects model is statistically significant at $p < 0.01$. Economically, an increase of one standard deviation in *RSI* is associated with an increase of yield spread of about 17 basis points.

The control variables across all models have their theoretically predicted signs and in general are statistically significant. We find firm size, profitability, growth opportunities, momentum, tangibility, and credit ratings to be negatively associated with yield spreads, while firm leverage and stock volatility are positively related to bond yield spreads. We also find that institutional ownership is negatively related to yield spreads, consistent with monitoring (Bhojraj and Sengupta 2003), and that liquidity is positively priced in the debt market as older bonds provide a higher yield spread than less liquid ones. Overall, the results indicate that short sellers are sophisticated investors that provide valuable information to the bond market.

V. ESTABLISHING CAUSALITY

We next examine whether the relation between short sales and the yield spread is causal. To establish causality, we conduct two tests. First, we follow Pownall and Simko (2005) and employ a short-window event study design around short interest spikes. We identify large increases and decreases in short interest and observe short-window bond returns that are significant around public disclosures of the short interest spikes. Second, we conduct a natural experiment using the SEC's Regulation SHO pilot program, which was designed to eliminate a constraint on short selling. We explain these two tests below.

Event Study Using Short Interest Spikes

We test for causality using an event study methodology as in Pownall and Simko (2005), who argue that short sellers provide valuable information to investors. We posit that the announcement of short interest is also informative to bond investors. Bond investors will view a positive (negative) short interest spike as providing new and negative (positive) information about the financial health of the firm and thus the default risk of the bond. Therefore, the bond market should react negatively (positively) around the public disclosure of positive (negative) spikes. Accordingly, we predict a significant negative (positive) abnormal bond return during the announcement period for firms that experience a positive (negative) spike in short interest.

We test this prediction by first constructing a measure of short interest spikes. We follow Pownall and Simko (2005) and use a standard market model approach based on a firm's observable change in short interest. Specifically, for each firm we compute the time-series of relative short interest (*RSI*). We then difference this series to calculate the change in short interest for firm i in month t (ΔRSI), and regress this change against the contemporaneous market average change in month t. The residuals from these models serve as our proxies for abnormal changes in short interest (ΔASI). That is:

$$\Delta RSI_{it} = \gamma_0 + \gamma_1 \Delta RSI_{mt} + \varepsilon_{it}, \tag{3}$$

$$\text{where } \varepsilon_{it} = \Delta ASI_{it}, \tag{4}$$

and we standardize the abnormal short interest change in month t (defined as $\Delta SASI_{it}$) by dividing

ΔASI_{it} by its standard deviation. We define positive and negative short interest spikes separately. For positive spike events, we follow Pownall and Simko (2005) and identify all firm-month observations meeting two additional criteria: (1) $\Delta SASI_{it}$ is 2.0 or greater in month t but is less than 1.2 in month t–1 (i.e., the initial positive spike), and (2) within the prior 12 months the positive short spike is not preceded by another that had not yet reverted back to its pre-spike short interest level. For this second criterion, we define a positive spike in month t as having reverted if the cumulative decline in short interest since the short spike is at least 75 percent of the initial increase in short interest during the short spike month t. Similarly, a firm in month t is identified as a negative spike event if the observation meets the above two criteria (as in Pownall and Simko [2005]), but in the opposite direction.

To determine whether a firm in month t has experienced a positive or negative spike, we use all observations (with a minimum of 24 months) up to month t to estimate the abnormal changes in short interest (Equation (3)) and obtain $\Delta SASI_{it}$. We impose a number of other restrictions to reach the final sample. First, since the TRACE bond daily dataset starts in July 2002, we only include spike events from that date to June 2011. Second, we exclude firms not included on both the CRSP and Compustat databases because our main yield spread test requires both stock and accounting information. Third, we exclude financial firms to avoid unnecessary complexities involving accounting rules and regulations of financial firms.[15] Fourth, we require the firm to have bond pricing information in the event window to compute bond returns.[16] Imposing the fourth requirement substantially reduces the sample size, an issue that we address below.

We hand-collect the short interest disclosure dates as reported in the *Wall Street Journal* and on the NYSE website. Major exchange members report their short positions as of the 15th of each month to the exchange, which then compiles the data and subsequently discloses the short interest to the public around the 21st (NYSE and AMEX) or 26th (NASDAQ). We thus have 108 disclosure dates for our sample period. We compute bond returns during the announcement period from business day $-l$ to business day $+m$ (i.e., $l+m+1$ day return) as:

$$Ret_{-l,+m} = \frac{P_m + AI_m - (P_{-l-1} + AI_{-l-1}) + C}{P_{-l-1} + AI_{-l-1}}, \tag{5}$$

where P is the price at the beginning and ending of the test window, with day 0 the disclosure date, AI is accrued coupon interest as of the trading day, and C is the coupon payment received during the event period. Because TRACE sometimes reports multiple trades for a particular bond at a given day, we use two methodologies to get the price for that bond on that date. The first method mimics the convention in the stock market, which uses the price for the last trade on that day as the price for that bond. The second method averages all bond prices on that day based on their trading volume. Both methods yield similar results, so we report results based on the first method for brevity.

We allow l to be a positive number for a likely information leakage. When l is 0, we are not taking into consideration any of the information leakage and thus only consider bond market

[15] Estimating Equation (3) using the full sample data on NYSE industrial firms from 1989 to 1998, we find 1,323 (unique) firms and 2,328 positive spike events (or firm-month observations) after applying these restrictions. These numbers are comparable to those reported in Table 1 of Pownall and Simko (2005) for their primary sample (1,333 firms and 2,301 events).

[16] For example, applying these restrictions to a six-day event window (0, +5), we find 340 unique firms or 591 positive spike events, of which 2 firm-months are in 2002, 21 in 2003, 41 in 2004, 86 in 2005, 75 in 2006, 86 in 2007, 82 in 2008, 89 in 2009, 86 in 2010, and 20 in 2011. Similarly, we find 360 unique firms or 671 negative spike events, of which 8 firm-months are in 2002, 32 in 2003, 39 in 2004, 89 in 2005, 94 in 2006, 92 in 2007, 104 in 2008, 91 in 2009, 86 in 2010, and 36 in 2011. Both positive and negative spike events are well distributed across months, with March having the greatest number of positive spike events (67) and May having the greatest number of negative spike events (81). All other calendar months have about 50 spike events.

TABLE 7

Short Interest Spikes and Abnormal Bond Returns

	Negative Spikes				Positive Spikes			
	Obs. (2)				Obs. (5)			
Window (1)	Bond	Firm	Abnormal Bond Return (3)	t-Statistic (4)	Bond	Firm	Abnormal Bond Return (6)	t-statistic (7)
(0,+1)	1,570	674	0.225**	(2.31)	1,350	591	−0.148**	(−2.17)
(0,+2)	1,525	674	0.237**	(2.00)	1,394	619	−0.017	(−0.22)
(0,+3)	1,538	685	0.266**	(2.30)	1,408	620	−0.269***	(−3.52)
(0,+4)	1,556	682	0.366**	(2.30)	1,275	572	−0.230***	(−2.42)
(0,+5)	1,532	671	0.371***	(3.36)	1,331	591	−0.169**	(−1.98)
(−1,+1)	1,599	692	0.141	(1.43)	1,386	585	−0.090	(−1.24)
(−1,+2)	1,583	691	0.360***	(2.87)	1,435	613	0.006	(0.08)
(−1,+3)	1,605	709	0.465***	(3.22)	1,447	622	−0.145*	(−1.91)
(−1,+4)	1,587	682	0.449**	(2.14)	1,331	591	−0.181*	(−1.92)
(−1,+5)	1,529	679	0.435***	(3.51)	1,378	590	−0.202**	(−2.29)

*, **, *** Significance at the 10 percent, 5 percent, and 1 percent levels, respectively.
This table provides results on buy-and-hold benchmark adjusted returns during announcement period from day l to day m, with l from −1 to 0 and m from +1 to +5. The benchmark returns are Merrill Lynch bond index returns matched by rating and maturity. Day 0 is the spike disclosure day hand-collected from the *Wall Street Journal* and exchange website. Bond pricing data are obtained from TRACE, with the sample period spanning from the period July 2002 to June 2011. We apply a market model to short interest data for NYSE non-financial firms on a rolling-window basis to determine the months when each individual firm experiences a short interest spike. We compute abnormal bond returns for firms using negative and positive spikes.

reaction from the announcement date and thereafter. Note that unlike stocks, bonds are infrequently traded and therefore not all bonds issued by the event firm have bond prices available at day $+m$ and day $(-l-1)$.[17] To obtain robust results, we consider l to be 1 and 0 and m to be any number from 0 to 5. For example, when $l=1$ and $m=5$, we open an event study window of seven days from −1 to +5.

After computing the bond return during the announcement period, we compound the Merrill Lynch daily bond index in the announcement window. We match each bond to the index by rating and maturity and adjust the raw bond return by the bond index return to obtain the abnormal bond return. Therefore, the calculated abnormal bond return has purged out the market effect as well as rating and maturity effects, such that our inference on the bond market reaction to the disclosure of short interest position is not subject to these compounding effects.

To test our prediction, we compute abnormal bond returns for negative and positive spike firms and report the results in Table 7. We also report the number of bond-months and firm-months. The two-day window from 0 to 1 yields 674 negative spikes for 1,570 bonds with valid data to compute abnormal bond returns. Bonds of negative spike firms have abnormal positive returns of about 22.5 bps, results that are statistically significant at $p < 0.05$. The same event window yields 591 positive spike events for 1,350 bonds that earn abnormal negative returns of 14.8 basis points (statistically significant at $p < 0.05$). When we extend the window from 0 to +5, we observe larger abnormal

[17] Because bonds do not trade continuously, abnormal returns may not be cumulative for all firms. For example, the abnormal return from 0 to 5 is not the sum of abnormal returns from 0 to 1 and from 1 to 5 because the sample bonds with valid bond trading information in these three days are not the same.

returns with the exception of $m = +2$ for positive spikes. We next allow a possible information leakage by opening the event window one day before the announcement date and find similar and significant negative abnormal return, suggesting that the impact of a potential information leakage is not material for the bond market.

There are two potential concerns in our analysis. Because we require firms in the sample to have valid pricing information in the event window to compute bond returns, about 60 percent of spike firms do not see their bonds traded in the event window. Therefore, these firms are not included in the final sample. This raises a question of whether the bond market is not responsive to the disclosure of short interest spike. We find that the number of firms with bond trades during the event period is, in general, similar to or greater than the one during the pre-event period. For example, the 31-day window from −30 to −1 yields 555 positive spike events for 1,257 bonds. Therefore, the fact that not all spike firms have bond trades in the event period is more likely due to the illiquidity in the bond market.

To ensure that the abnormal negative bond returns around the disclosure of positive spikes is not a continuation of prior bond price movement, we performed the analysis using a larger window of (−30, −1). Specifically, we open up the event window to various periods prior to the announcement day for positive spikes (i.e., from −30 to −1) with increments of 5 (i.e., −25, −20, −15, −10, and −5 days to 1 business day before the announcement date) and find insignificant results (not tabulated). Therefore, the abnormal bond returns during the event window period are less likely to be a continuation of prior bond price movement. We also conduct a multivariate regression analysis on event period abnormal bond returns, where we control for firm size, stock return momentum, and other firm- and security-specific characteristics. The estimated coefficient (not tabulated) on prior bond return is significant and negative, evidence that is inconsistent with the view that the event period bond return is just caused by bond return momentum. Overall, the evidence is consistent with our prediction that the bond market perceives a positive (negative) spike in short interest as a negative (positive) event and thus short sellers provide valuable information to bond investors.

Natural Experiment

Description of the Regulation SHO Pilot Program

We use the SEC's Regulation SHO pilot program to identify changes in short sales that are exogenous to changes in the cost of debt. Adopted in the Securities Exchange Act of 1934 and made effective in 1938, the "uptick rule" (Rule 10a-1) requires that every short sale occur at a price that is higher than the price of the previous trade. The consensus is that the uptick rule was a binding short sales constraint (e.g., Alexander and Peterson 1999). In 2004, the SEC announced that it would conduct a year-long experiment—a "pilot program"—to study the effect of the uptick rule on market quality.

The experiment involved temporarily suspending the uptick rule for a random sample of publicly traded firms and studying its effect by comparing this sample of firms to other publicly traded firms for which the uptick rule would remain effective. The timeline for the pilot program is as follows: on June 23, 2004, the SEC adopted the pilot program; on June 25, the SEC selected the list of stocks that would be included in the pilot program; and on July 28, the SEC announced the details of the experiment including the list of stocks in the program and the dates of the program (May 2, 2005 to April 28, 2006).

To select the stocks in the pilot program, the SEC began with all stocks in the Russell 3000 index, which consists of the largest 3,000 U.S. stocks as of the end of May 2004. Within each of the three stock exchanges (NYSE, NASDAQ, and AMEX), the SEC ranked all of the stocks in the Russell 3000 index according to their average daily trading volume, and picked every third stock

for inclusion in the pilot program (i.e., for temporarily suspending the uptick rule). This sample forms the basis of our "treatment firms," although there were actually 986 firms in the pilot program. The other two-thirds form the basis of our "control firms."

As a result, the Regulation SHO pilot program provides us with a natural experiment in which a random sample of the biggest publicly traded firms had their short sales constraints relaxed, while the rest of these firms had their short sales constraints maintained. As expected, the empirical evidence shows short sales significantly increased after the uptick rule was suspended for stocks in the pilot program compared to those not in the pilot program (e.g., see SEC 2007; Alexander and Peterson 2008; Boehmer et al. 2008; Diether et al. 2009). We use this exogenous shock to test whether changes in short sales cause changes in the cost of debt.

Data and Sample

We begin our analysis by determining the appropriate event date. We use June 25, 2004 as the date at which the stocks were selected for the pilot program. The rationale is that stock prices appear to impound the information about the pilot program list beginning in mid-July 2004 and the market reaction is largely complete (roughly −2 percent to reflect the relaxation of short sales constraints) by the announcement date of July 28, 2004 (e.g., Grullon et al. 2011). Thus, bond prices should similarly react long before the start date of the program of May 2, 2005. Since our data on the cost of debt after the shock are from 2005 (see below), the cost of debt will have started to reflect the shock (because almost 80 percent of our sample firms have fiscal years that end in December). Therefore, even if short sales do not affect the cost of debt until the start date of the pilot program, our timing will still capture the effect, albeit with more noise.

To construct a sample of "treatment firms" and "control firms," we follow Grullon et al. (2011) and construct a list of U.S. stocks included in the Russell 3000 index at the end of May 2004. This is the list of stocks that the SEC selected for inclusion in the pilot program on June 25, 2004. We obtain the list of the stocks in the pilot program from the SEC. To perform our analysis, we use a difference-in-differences approach by comparing the treatment firms before and after the shock to the control firms before and after the shock. Since the treatment firms and control firms are necessarily similar because their selection is random, our approach ensures that changes in the cost of debt are caused by changes in short sales. Accordingly, we do not control for cross-sectional and time-series effects that affect both short sales and the cost of debt.

Due to data constraints, we do not use the bond market cost of debt from the Lehman Brothers Fixed Income database because it is only available for roughly 20 percent of the Russell 3000 index. By contrast, accounting data are available for almost 90 percent of the data, and the accounting cost of debt is available for 1,580 firms or more than 50 percent of the Russell 3000. The difference arises because, for a firm to be included in our sample, we require it to have at least \$100 million of long-term debt to be comparable to the bond market cost of debt.[18] Our sample comprises 1,580 firms, of which 542 are treatment firms and 1,038 are control firms.

We follow the literature and measure the annual cost of debt as interest expense divided by long-term debt (including current and non-current portions), where interest expense is computed during the year while long-term debt is computed as the average of long-term debt at the end of the previous year and at the end of the current year (e.g., Kim et al. 2011; Minnis 2011; Francis et al. 2005a; Francis et al. 2005b; Pittman and Fortin 2004). We measure the cost of debt before the shock using data from the fiscal year end date during 2003, and after the shock using data from the

[18] We impose the \$100 million long-term debt requirement because only 1 percent of firms with bond market data have long-term debt of less than \$100 million. Therefore, this requirement ensures that the samples are similar, but it does not materially affect our results.

fiscal year end date during 2005. This ensures that the accounting cost of debt captures the changes in the cost of debt caused by the shock.

To validate our measure of the cost of debt, we examine the correlation between yields from bond market data and accounting data. During the year before the shock, the correlation is 0.530 ($p < 0.01$), and during the year after the shock, the correlation is 0.459 ($p < 0.01$). This suggests that the accounting data are similar to the bond market data, such that it is reasonable for us to use them to study a more comprehensive sample of firms in the Russell 3000 index.

Results

As would be expected under random selection, Table 8, Panel A shows that treatment firms are similar to control firms during the year before the shock in terms of their cost of debt, amount of debt, market capitalization, and market-to-book values. Additionally, a similar proportion of them are financially constrained (not tabulated).

Next, we examine the causal effect of short sales on the cost of debt using the shock to short sales constraints and thus to short sales themselves. We find that, on average, the cost of debt changes by 7.4 basis points (8.7 – 1.3) for treatment firms compared to control firms, although the increase is not statistically significant. An economically small effect is not unexpected, given that our sample firms are, by construction, the biggest publicly traded firms, even bigger than the average firm in the Russell 3000 index because of our requirement that each sample firm has at least $100 million of long-term debt.

We also examine whether the effect of short sales on the cost of debt financing is larger for smaller firms as well as firms that are financially constrained (i.e., those that do not generate sufficient internal financing and have limited access to the capital market). If the relaxation of short sale constraints increases short selling in the equity market, this will cause more release of negative information about the firm. To the extent that equity short selling activity provides a signal about the credibility and quality of financial reporting, we expect the effect of equity short interest on the cost of debt to be higher for financially constrained firms.

We find evidence consistent with our conjecture. Among small and financially constrained firms, the increase in the cost of debt financing for treatment firms is significantly higher than that for control firms (25 basis points). We test whether the economically significant increases in the cost of debt for small and financially constrained firms are statistically significant. Table 8, Panel B shows that these increases are generally statistically significant. Moreover, the changes in the cost of debt for financially unconstrained large firms are not statistically significant. Indeed, the triple differences are generally even more statistically significant than the double differences for small and financially constrained firms. Our results are also significant in absolute magnitude: the median and mean firms in our sample have long-term debt of about $600 million and $5 billion, respectively, so an increase in the cost of debt of 25 basis points translates into an annual interest expense of $1.5 million and $12.5 million, respectively.

Figure 1 splits the unconditional difference-in-differences in the cost of debt by market capitalization and proxies for financial constraints (K-Z index, dividend payer status, bond rating status, and the cash flow-investment gap). The change in the cost of debt is economically significant for small firms and firms that are financially constrained, ranging from an increase of 18 basis points to an increase of 49 basis points. By contrast, the change in the cost of debt is economically insignificant for big firms and firms that are not financially constrained.

In summary, using the natural experiment of the SEC's Regulation SHO pilot program, we find that an increase in short sales causes an economically small increase in the cost of debt. This increase becomes economically and statistically significant for small firms and firms that are

TABLE 8

Effect of Short Sales on the Cost of Debt: Evidence from a Natural Experiment

Panel A: Descriptive Statistics

	Mean		25th Percentile		Median		75th Percentile		p-value of Test of Equality of Distributions
	Treatment Firms	Control Firms	Treatment Firms	Control Firms	Treatment Firms	Control Firms	Treatment Firms	Control Firms	
Cost of Debt (bps)	544	540	390	340	583	590	742	744	0.735
Amount of Debt ($M)	2,468	3,301	238	253	574	600	1,685	1,682	0.865
Market Cap. ($M)	5,747	6,173	642	644	1,522	1,594	4,336	4,142	0.714
Market-to-Book	2.38	2.45	1.30	1.36	1.89	1.92	2.80	2.87	0.594

Panel B: Results from a Natural Experiment

Mean Cost of Debt (bps)	Market Capitalization	Kaplan-Zingales (K-Z) Index	Dividend Payer Status	Bond Rating	Cash Flow-Investment Gap
Financially Distressed Firms					
$\Delta T - \Delta C$ (1)	Small (n = 790)	High K-Z Index (n = 723)	Non-Payer (n = 646)	Non-Inv. Grade (n = 439)	Low Gap (n = 408)
Change in Treatment (ΔT)	7.2	−9.7	−11.3	−27.8	14.3
Change in Control (ΔC)	−12.5	−27.6	−40.1	−51.4	−34.7
Diff-in-Diff ($\Delta T - \Delta C$)	19.7*	17.9*	28.8**	23.6	48.9***
t-statistic	1.88	1.96	2.12	1.62	3.18
Financially Non-Distressed Firms					
$\Delta T - \Delta C$ (2)	Big (n = 790)	Low K-Z Index (n = 724)	Payer (n = 934)	Inv. Grade (n = 614)	High Gap (n = 408)
Change in Treatment (ΔT)	10.4	30.6	23.7	32.1	17
Change in Control (ΔC)	14.7	33	28.5	26	23.1
Diff-in-Diff ($\Delta T - \Delta C$)	−4.3	−2.4	−4.8	6	−6.1
t-statistic	−0.48	−0.2	−0.71	0.78	−0.35

(continued on next page)

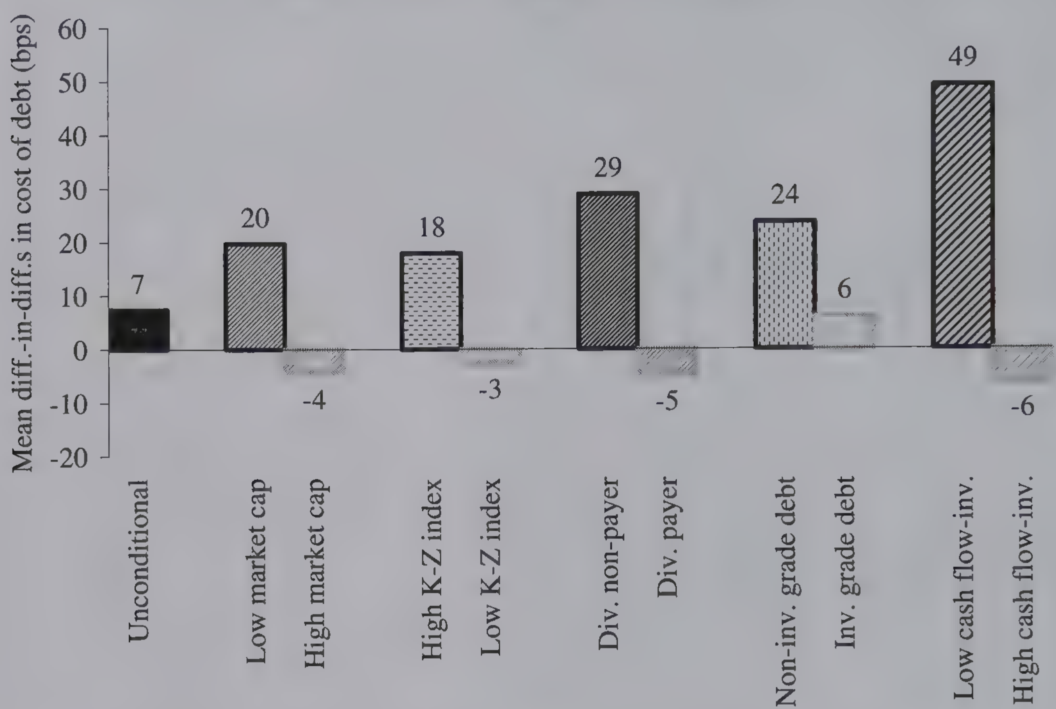
TABLE 8 (continued)

Mean Cost of Debt (bps)	Market Capitalization	Kaplan-Zingales (K-Z) Index	Dividend Payer Status	Bond Rating	Cash Flow-Investment Gap
Financially Distressed Firms – Financially Non-Distressed Firms					
$\Delta T - \Delta C$ (1) – $\Delta T - \Delta C$ (2)	Small – Big	Low – High	Non-Payer – Payer	Non-Inv. Grade – Inv. Grade	Low Gap – High Gap
Mean	24.0**	20.3*	33.7***	17.5*	55.0***
t-statistic	2.47	1.90	3.59	1.68	3.36

*, **, *** Significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

Panel A provides descriptive statistics for the natural experiment, and Panels B, C, and D present the results from the experiment. The sample comprises 1,580 firms, of which 542 are treatment firms and 1,038 are control firms. All descriptive statistics are measured before the natural experiment. The natural experiment is the temporary suspension of the uptick rule under the SEC's Regulation SHO pilot program. A random sample of the biggest publicly traded firms had their short sales constraints relaxed while the rest of these firms had their short sales constraints maintained. The cost of debt is measured as interest expense divided by long-term debt as in Pittman and Fortin (2004). The change in the cost of debt is measured from the year before the natural experiment to the year after the natural experiment. Financial constraints are proxied by the Kaplan-Zingales (K-Z) index of financial constraints, dividend payer status, bond rating status, and the cash flow-investment gap. Market capitalization and financial constraints are measured before the shock. The K-Z index (Kaplan and Zingales 1997) is computed using a weighted average of cash flows, market-to-book, leverage, dividends, and cash holdings. We classify firms in the top and bottom halves of the K-Z index as constrained and unconstrained, respectively. We classify dividend payers and firms with investment-grade bonds as non-constrained and dividend non-payers and firms with non-investment-grade bonds as constrained. The cash flow-investment gap is defined as operating cash flow divided by capital expenditures minus one. We classify firms in the bottom and top percentiles of the gap as constrained and unconstrained, respectively.

FIGURE 1
The Effect of Short Sales on the Cost of Debt Unconditionally and Conditionally Upon Size and Proxies for Financial Constraints



The cost of debt is measured as the mean difference-in-differences: (treatment firms after minus treatment firms before) minus (control firms after minus control firms before). The sample comprises 1,580 firms, of which 542 are treatment firms and 1,038 are control firms.

financially constrained. Overall, our results suggest that changes in short sales in the stock market cause changes in yield spreads in the bond market.

VI. CONCLUSION

We examine the relation between short interest and the cost of debt financing. We find a statistically and economically positive and significant relation between equity short interest and bond yield spreads after controlling for firm- and security-specific factors. We also find that the relation between equity short interest and yield spreads is affected by the information environment of the firm. The predictability and informativeness of equity short interest for bond pricings decrease in the quantity and quality of information provided by equity analysts. Accordingly, less analyst following and high forecast dispersion will enhance the equity short interest effect on the cost of debt financing. We also document that equity short interest can predict various credit downgrades in the next month. Overall, the evidence suggests that short selling activities in the stock market provide value relevant information to investors in the bond market, and firms with high levels of equity short interest are viewed unfavorably in the debt market.

One of the main concerns of this study is whether the relation between short interests and bond spreads is causal or merely associative. Accordingly, we conduct two distinct tests. First, we follow Pownall and Simko (2005) and employ a short-window event study design around positive as well as negative short interest spikes. We document significant abnormal bond returns of −17 basis points for the positive spikes and +37 basis points for the negative spikes in the event window (0,+5). Second, we conduct a natural experiment employing the SEC's Regulation SHO pilot

program. Using this exogenous shock we find that an increase in short sales causes an increase in the cost of debt. The results are more pronounced in small firms and firms that are financially constrained.

Because bond investors are overwhelmingly sophisticated institutional investors, our evidence indicates that equity short sellers can provide useful information for bond investors to price debt and evaluate firm credit quality. Our finding that equity short selling conveys significant external benefits to bond investors suggests that any legal restriction on equity short selling runs the risk of limiting a potentially important source of information for investors in the corporate debt market. This finding also suggests that it is important to look at the total effect of any restrictions on equity short selling on all capital markets and not merely the impact on stockholders before drawing conclusions about regulatory policy.

REFERENCES

Alexander, G., and M. Peterson. 1999. Short selling on the New York Stock Exchange and the effects of the uptick rule. *Journal of Financial Intermediation* 8 (1-2): 90–116.

Alexander, G., and M. Peterson. 2008. The effect of price tests on trader behavior and market quality: An analysis of Reg SHO. *Journal of Financial Markets* 11 (1): 84–111.

Anderson, R., S. Mansi, and D. Reeb. 2003. Founding family ownership and the agency costs of debt. *Journal of Financial Economics* 68 (2): 263–285.

Anderson, R., S. Mansi, and D. Reeb. 2004. Board characteristics, accounting report integrity, and the cost of debt. *Journal of Accounting and Economics* 37 (3): 315–342.

Asquith, P., P. Pathak, and J. Ritter. 2005. Short interest, institutional ownership, and stock returns. *Journal of Financial Economics* 78 (2): 243–276.

Beaver, W. 1998. *Financial Reporting: An Accounting Revolution*. Upper Saddle River, NJ: Prentice Hall.

Beim, D. 1992. *Estimating Bond Liquidity*. Working paper, Columbia University.

Bessembinder, H., W. Maxwell, and K. Venkataraman. 2005. Market transparency, liquidity externalities, and institutional trading costs in corporate bonds. *Journal of Financial Economics* 82 (2): 251–288.

Bharath, S., and T. Shumway. 2008. Forecasting default with the KMV-Merton model. *Review of Financial Studies* 21 (3): 1339–1369.

Bhojraj, S., and P. Sengupta. 2003. The effect of corporate governance mechanisms on bond ratings and yields: The role of institutional investors and outside directors. *Journal of Business* 76 (3): 455–475.

Boehmer, E., C. Jones, and X. Zhang. 2008. Which shorts are informed? *Journal of Finance* 63 (2): 491–527.

Campbell, J., and G. Taksler. 2003. Equity volatility and corporate bond yields. *Journal of Finance* 58 (6): 2321–2349.

Chanos, J. 2003. *Prepared Statements*. U.S. Securities and Exchange Commission Roundtable on Hedge Funds, May 15.

Chen, X., Q. Cheng, A. K. Lo. 2012. Accounting restatements and external financial choices. *Contemporary Accounting Review* (forthcoming).

Christophe, S. E., M. G. Ferri, and J. J. Angel. 2004. Short selling prior to earnings announcements. *Journal of Finance* 59: 1845–1875.

Christophe, S., M. Ferri, and J. Hsieh. 2010. Informed trading before analyst downgrades: Evidence from short sellers. *Journal of Financial Economics* 95 (1): 85–106.

Coles, J., M. Lemmon, and F. Meschke. 2012. Structural models and endogeneity in corporate finance. *Journal of Financial Economics* 103 (1): 149–168.

Dechow, P., A. Hutton, L. Meulbroek, and R. Sloan. 2001. Short-sellers, fundamental analysis, and stock returns. *Journal of Financial Economics* 61 (1): 77–106.

DeFond, M., and J. Jiambalvo. 1994. Debt covenant violation and manipulation of accruals. *Journal of Accounting and Economics* 17 (1-2): 145–176.

Desai, H., S. Krishnamurthy, and K. Venkataraman. 2006. Do short sellers target firms with poor earnings quality? Evidence from earnings restatements. *Review of Accounting Studies* 11 (1): 71–90.

Diamond, D., and R. Verrecchia. 1987. Constraints on short selling and asset price adjustment to private information. *Journal of Financial Economics* 18 (2): 277–311.

Dichev, I., and J. Piotroski. 2001. The long-run stock returns following bond ratings changes. *Journal of Finance* 56 (1): 173–203.

Diether, K., K. Lee, and I. Werner. 2009. It's SHO time! Short-sale price tests and market quality. *Journal of Finance* 64 (1): 37–73.

Downing, C., S. Underwood, and Y. Xing. 2009. An intraday analysis of the relative informational efficiency of stocks and bonds. *Journal of Financial and Quantitative Analysis* 44 (5): 1081–1102.

Drake, M., L. Rees, and E. Swanson. 2011. Should investors follow the prophets or the bears? Evidence on the use of public information by analysts and short sellers. *The Accounting Review* 86 (1): 101–130.

Edwards, A., L. Harris, and M. Piwowar. 2007. Corporate bond market transparency and transaction costs. *Journal of Finance* 62 (3): 1421–1451.

Efendi, J., and E. Swanson. 2009. *Short Seller Trading in Companies with Accounting Irregularities*. Working paper, The University of Texas at Arlington.

Einhorn, D. 2008. *Fooling Some of the People All of the Time*. Hoboken, NJ: Wiley.

Fleming, J., C. Kirby, and B. Ostdiek. 1998. Information and volatility linkages in the stock, bond, and money markets. *Journal of Financial Economics* 49 (1): 111–137.

Francis, J., I. Khurana, and R. Pereira. 2005a. Disclosure incentives and effects on cost of capital around the world. *The Accounting Review* 80 (4): 1125–1162.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2005b. The market pricing of accruals quality. *Journal of Accounting and Economics* 39 (2): 295–327.

Francis, J., M. Venkatachalam, and Y. Zhang. 2006. *Do Short Sellers Convey Information about Changes in Fundamentals or Risk?* Working paper, Duke University.

Gebhardt, W., S. Hvidkjaer, and B. Swaminathan. 2005. Stock and bond market interaction: Does momentum spill over? *Journal of Financial Economics* 75 (3): 651–690.

Gow, I., G. Ormazabal, and D. Taylor. 2010. Correcting for cross sectional and time series dependence in accounting research. *The Accounting Review* 85 (2): 483–512.

Graham, J., S. Li, and J. Qiu. 2008. Corporate misreporting and bank loan contracting. *Journal of Financial Economics* 88 (1): 44–61.

Grullon, G., S. Michenaud, and J. Weston. 2011. *The Real Effects of Stock Market Prices*. Working paper, Rice University.

Henry, T., D. Kisgen, and J. Wu. 2010. *Do Equity Short Sellers Anticipate Bond Rating Downgrade?* Working paper, The University of Georgia.

Hirshleifer, D., S. Teoh, and J. Yu. 2011. Short arbitrage, return asymmetry and the accrual anomaly. *Review of Financial Studies* 24 (7): 2429–2461.

Hotchkiss, E., and T. Ronen. 2002. The informational efficiency for the corporate bond market: An intraday analysis. *Review of Financial Studies* 15 (5): 1325–1354.

Jorion, P., Z. Liu, and C. Shi. 2005. Informational effects of Regulation FD: Evidence from rating agencies. *Journal of Financial Economics* 76 (2): 309–330.

Kaplan, S., and L. Zingales. 1997. Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics* 112 (1): 169–215.

Karpoff, J., and X. Lou. 2010. Short sellers and financial misconduct. *Journal of Finance* 65 (5): 1879–1913.

Kim, J., D. Simunic, M. Stein, and C. Yi. 2011. Voluntary audit and the cost of debt capital for privately held firms: Korean evidence. *Contemporary Accounting Research* 28 (2): 585–615.

Kwan, S. 1996. Firm-specific information and the correlation between individual stocks and bonds. *Journal of Financial Economics* 40 (1): 63–80.

Mansi, S., W. Maxwell, and D. Miller. 2004. Does auditor quality and tenure matter to investors? Evidence from the bond market. *Journal of Accounting Research* 42 (4): 755–793.

Mansi, S., W. Maxwell, and D. Miller. 2011. Analyst forecast characteristics and the cost of debt. *Review of Accounting Studies* 16 (1): 116–142.
McConnell, J., and H. Servaes. 1990. Additional evidence on equity ownership and corporate value. *Journal of Financial Economics* 27 (2): 595–612.
Minnis, M. 2011. The value of financial statement verification in debt financing: Evidence from private U.S. firms. *Journal of Accounting Research* 49 (2): 457–506.
Morck, R., A. Shleifer, and R. Vishny. 1990. The stock market and investments: Is the market a side-show? *Brookings Papers on Economic Activity* 2: 157–215.
Norden, L., and M. Weber. 2004. Informational efficiency of credit default swap and stock markets: The impact of credit rating announcements. *Journal of Banking and Finance* 28 (11): 2813–2843.
Pittman, J., and S. Fortin. 2004. Auditor choice and the cost of debt capital for newly public firms. *Journal of Accounting and Economics* 37 (1): 113–136.
Pownall, G., and P. Simko. 2005. The information intermediary role of short sellers. *The Accounting Review* 80 (3): 941–966.
Qi, C., K. Subramanyam, and J. Zhang. 2010. *Accrual Quality, Bond Liquidity, and Cost of Debt.* Working paper, University of Southern California.
Schilit, H. 2002. *Financial Shenanigans*, 2nd edition. New York, NY: McGraw-Hill.
Securities and Exchange Commission (SEC). 2007. Economic analysis of the short sale price restrictions under the Regulation SHO pilot. Available at: http://www.sec.gov/news/studies/2007/regshopilot020607.pdf
Shi, C., and S. Zhang. 2008. *Accounting Restatements and the Cost of Debt Capital.* Working paper, University of California, Irvine.
Staley, K. 1997. *The Art of Short Selling*. New York: Wiley.

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 641–665

American Accounting Association
DOI: 10.2308/accr-50332

Financial Statement Disaggregation Decisions and Auditors' Tolerance for Misstatement

Robert Libby
Cornell University

Timothy Brown
University of Illinois at Urbana–Champaign

ABSTRACT: Current IFRS requires significant disaggregation of income statement numbers while such disaggregation is voluntary and much less common under U.S. GAAP. We examine whether voluntary disaggregation of income statement numbers increases the reliability of income statement subtotals because auditors permit less misstatement in the disaggregated numbers. In our experiment, experienced auditors require correction of smaller errors in disaggregated numbers. Auditors also believe that greater disaggregation will increase SEC scrutiny of uncorrected financial statement errors in the disaggregated numbers. However, the effects are substantially reduced if the disaggregated numbers are presented in the notes. Furthermore, there is significant disagreement among participants on whether disaggregated numbers are relevant materiality benchmarks, and on what current auditing guidance requires. These results suggest a potential deficiency in current audit guidance, which traditionally has been aimed at promoting consensus in practice among auditors. The results also suggest an unintended positive consequence of voluntary disaggregation for the reliability of income statement subtotals. Possible effects of management behavior and required disaggregation resulting from U.S. adoption of IFRS or the recommendations of the joint FASB/IASB financial statement presentation project are also discussed.

Keywords: *disaggregation*; *materiality*; *IFRS vs. U.S. GAAP*; *statement recognition vs. note disclosure*; *audit guidance.*

Data Availability: *Data are available from the authors upon request.*

We thank Scott Asay, Scott Emett, Marsha Keune, Mark Nelson, Tammie Rech, Kristina Rennekamp, Scott Vandervelde, participants at workshops at The University of Chicago, Cornell University, Georgia Institute of Technology, Nanyang Business School, University of North Texas, University of South Carolina, Temple University, and Texas A&M University, and the anonymous reviewers and editors for their helpful comments. We also thank the auditors who participated in our experiments and the three participating Big 4 firms.

Editor's note: Accepted by Donald Moser.

Submitted: August 2011
Accepted: October 2012
Published Online: October 2012

I. INTRODUCTION

The joint FASB/IASB financial statement presentation project is considering the most dramatic changes to the format of, and information included in, the four basic financial statements and notes since the requirement to present a cash flow statement (see e.g., FASB 2010, July 1; hereafter, Staff Draft[1]). In a survey of analysts (FASB/IASB 2009, September 21; hereafter, the Analyst Survey), respondents judged increased disaggregation to be the most useful change in the proposed presentation (even ahead of the direct method cash flow statement). The related discussion documents also provide options for preparers to include disaggregated information on the face of the statements or in the notes (see e.g., FASB 2008, October 16; hereafter, Preliminary Views).

While the outcome of the joint FASB and IASB deliberations is still uncertain, substantial differences in the level of disaggregation of financial statement numbers already exist between U.S. GAAP and IFRS companies (following Regulation S-X and IAS 1, respectively) and, to a lesser but significant extent, among companies applying each set of standards (see e.g., SEC 2011a). For example, like most U.S. retailers, Wal-Mart splits total operating expenses into only two categories: cost of sales and selling, general, and administrative (SG&A). Furthermore, its additional note disclosures explain the nature of less than 5 percent of the total SG&A expenses. In contrast, the U.K. retailer Marks & Spencer breaks down SG&A into selling and marketing versus administrative expenses on the face of the income statement and further disaggregates each of these two categories into six subcategories in the notes. The Marks & Spencer statement and note presentation closely mimics the income statement disaggregation suggested in the Staff Draft.[2] The Analyst Survey results suggest that this type of disaggregation meets the goal of enhancing the usefulness of the information in predicting the entity's cash flows.

In this study, we investigate the possibility that disaggregation as well as the location of disaggregated numbers will affect the behavior of auditors in a manner that also affects the reliability of bottom line net income and other statement subtotals and totals. Auditors require managers to correct discovered financial statement errors only when they are deemed material, and recent evidence suggests that reporting managers leave many immaterial errors uncorrected (Keune and Johnstone 2009, 2012). Net income (or net income before taxes) is the most often cited quantitative materiality benchmark. Disaggregation can affect the reliability of income statement subtotals and totals if auditors judge the *smaller* disaggregated numbers, in addition to subtotals and totals, to be relevant reporting materiality benchmarks.

We examine (1) whether disaggregation increases the reliability of income statement numbers by decreasing the amount of misstatement that auditors tolerate, (2) whether location (statement or note presentation) of the disaggregated data moderates this effect, (3) the degree of consensus among auditors on these two issues, and (4) what concerns drive auditors' consideration of disaggregated line item amounts in judging the materiality of discovered financial statement errors (e.g., changed expectations about the impact on informed users and/or possible actions of regulators). The answers to these questions are relevant to financial reporting standard-setters and regulators who wish to consider the indirect effects of their financial statement presentation standards on the reliability of the information presented, to auditing standard-setters and regulators who wish to clarify auditors' responsibility for misstatement in disaggregated numbers, and to audit firms that provide guidance to ensure consensus in their auditors' judgments in light of changes in

---

[1] The most recent Staff Draft of an Exposure Draft on Financial Statement Presentation is dated July 1, 2010.
[2] The Preliminary Views note that companies following IFRS should already be disaggregating financial statement information in this fashion.

statement aggregation and format. As we discuss in more detail in the next section, existing professional guidance to auditors on these issues is unclear.

In our experiment, experienced audit managers evaluate a $79 million underaccrual of occupancy costs (part of selling and marketing expenses) discovered late in the audit of a large retailer. Importantly, all participants have access to disaggregated trial balance data. The auditors are assigned to one of three public reporting conditions: aggregated, disaggregated-face, and disaggregated-note. The $79 million error is approximately 4.5 percent of pretax income, 3.8 percent of selling and marketing expenses (disclosed to the public in all conditions), and 18.1 percent of occupancy costs (disclosed to the public only in the disaggregated conditions). Participants are told this is the only discovered audit difference and that the client objects to any adjustments to the statements and disclosures. The participants answer whether the failure to correct the misstatement of occupancy costs would result in statements that are materially misstated and the maximum dollar understatement of the expense that would be immaterial. The participants then complete a series of questions including manipulation checks, within-subjects comparisons of an error in statements with differing levels of aggregation, and other questions designed to examine their underlying cognitive processes.

Our between-subjects results, which reflect participants' natural reasoning process, indicate that disaggregating expense items on the face of the income statement significantly decreases the amount of error that auditors will accept. Disaggregation on the face of the financial statements reduces allowable error by 17 percent on average, but increases the variance in responses significantly. Furthermore, the effect of disaggregation on allowable error is substantially reduced if the disaggregated expense items are instead included in the notes. This suggests that line items are relevant materiality benchmarks for some auditors, but there is little consensus on the issue. Moving the disaggregated data to the notes virtually eliminates their relevance as materiality benchmarks. The within-subjects results and answers to open-ended questions better reflect participants' beliefs or knowledge. These results show greater effects, but again little consensus, with 58 percent of participants indicating that disaggregating expenses will increase the materiality of the error, while 42 percent indicate otherwise. Again, the effect on materiality judgments decreases significantly when the disaggregated amounts are included in the notes. Answers to open-ended questions reflect widespread disagreement on the interpretation of admittedly confusing current auditing standards related to this issue.

An error of the same percentage magnitude as the error in our case would overstate earnings per share (EPS) by $0.02 per share for Marks & Spencer and $0.17 per share for Wal-Mart. Recent capital markets research indicates that EPS changes of these magnitudes have significant price effects (e.g., Bhojraj et al. 2009). This suggests that changes in audit decision making caused by voluntary disaggregation choices, as well as current differences between the disaggregation required by U.S. GAAP and IFRS and possible changes brought about by the FASB/IASB financial statement presentation project, could be economically significant to the capital markets. As a consequence, audit firms, auditing standard-setters, and regulators may want to clarify the implications of current audit guidance for this issue. More generally, the results suggest that subtle differences between alternative financial reporting standards can have unintended consequences for audit effectiveness and capital markets. In this case, the unintended effect should be to increase the reliability of the resulting financial statement subtotals by reducing auditors' materiality thresholds. But as our results for note presentation suggest, if financial reporting standard-setters choose to require or allow note presentation, then this could eliminate positive effects or result in negative effects (Libby et al. 2006).

Section II discusses relevant professional and research literature and develops our hypotheses. Section III describes our experimental design and method. Section IV discusses the results. Section V concludes.

II. RELEVANT LITERATURE AND HYPOTHESES

Disaggregating the Income Statement

Schipper (2007) describes disaggregation to aid prediction as one of three primary purposes of disclosure standards. She discusses disaggregation of revenues into same store and new store sales to aid in prediction of revenue and segment disclosures to aid in prediction of earnings as examples of currently required disaggregations that meet this goal. The joint FASB/IASB financial statement presentation project similarly motivates its recommendation to disaggregate income and expense by function and nature "to the extent that this will enhance the usefulness of information in predicting the entity's cash flows (Preliminary Views, para. 3.42)." Function refers to the primary activities in which an entity is engaged (for example, selling goods or administration). Nature refers to the economic characteristics or attributes that distinguish expense items that do not respond equally to similar economic events (for example, disaggregating selling expense into labor, occupancy, advertising, etc.).

IAS 1 (IASB 2009) currently requires disaggregation by function *or* nature, whichever is more relevant, but if management classifies expenses by function, then additional disaggregation by nature is required. The degree of disaggregation is left up to the company, and a recent SEC study (2011a) finds significant differences in income statement disaggregation under IAS 1. In contrast, U.S. GAAP (Regulations S-X and S-K) provides little current guidance concerning disaggregation of the income statement.

The potential dramatic differences in income statement disaggregation between that provided by some companies under IAS 1 (and recommended in the Staff Draft) and typical current U.S. practice are illustrated in Figure 1.

Figure 1, Panel A presents income statement data following the format used by most U.S. retailers such as Wal-Mart and Macy's. Note that selling, marketing, and administrative expense is a *single* line item. Panel B is prepared using the income statement disaggregation under IAS 1 and suggested in the Staff Draft. The categories used to disaggregate expense data and amounts are taken from the U.K. retailer Marks & Spencer. Note that selling, marketing, and administrative expense is broken down into *12* line items.[3] Also, the breakdown by nature separates expenses that respond differently to similar economic events: many employee costs are variable (sales staff are laid off as sales fall); depreciation and amortization are mostly fixed; and occupancy costs (rentals with sales contingencies) have characteristics of both. Bloomfield et al. (2011) demonstrate how such disaggregation can affect credit analysts' ability to identify a firm's operating cost structure.

Materiality of Discovered Misstatements

The probability of a material error in published financial statements is equal to the probability of an error in the unaudited numbers times 1 minus the probability that the error is discovered *and corrected*. The probability that a discovered error is corrected is a function of the size and importance of the error, or its materiality.[4]

---

[3] IAS 1 currently allows complete income statement disaggregation as presented by Marks & Spencer or partial disaggregation of important expenses as presented by a number of other U.K. retailers (IAS 1 as amended December 2009). The SEC (2011a) recently analyzed the application of IFRS by 183 companies from a variety of industries and countries and found significant variation across countries and industries. We further discuss these findings in Section V. The U.K. retail industry appears to follow the IFRS requirements quite closely. We also note that numerous U.S. retailers would include many occupancy costs and repairs and maintenance of property related to selling in cost of sales.

[4] Prior theory and research also indicates that auditors consider offsetting incentives to thwart or allow aggressive reporting that is often characterized as a trade-off between possible costs resulting from a lawsuit, regulatory action, and reputation loss versus the risk of client loss (e.g., Nelson 2004).

FIGURE 1
Typical U.S. GAAP Income Statement Presentation Compared to FASB/IASB Staff Draft Recommended Presentation

Panel A: Current U.S. GAAP

Net sales	$12,262.1
Cost of sales	(7,690.2)
Gross profit	4,571.9
Selling, marketing, and admin. expens	(2,644.5)
Operating profit	1,927.4
Interest, net	(164.5)
Income before income taxes	1,762.9
Income tax expense	(617.0)
Net income	$ 1,145.9

Panel B: Staff Draft Proposal

Net sales	$12,262.1
Cost of sales	(7,690.2)
Gross profit	4,571.9
Selling and marketing expenses	(2,074.4)
Administrative expenses	(570.1)
Operating profit	1,927.4
Interest, net	(164.5)
Income before income taxes	1,762.9
Income tax expense	(617.0)
Net income	$ 1,145.9
Notes to Financial Statements:	
Selling and marketing expenses include the following:	
Employee costs	(923.2)
Occupancy costs	(439.2)
Repairs & maint. of property	(76.6)
Depreciation	(343.5)
Amortization	(24.6)
Other costs	(267.3)
	$ (2,074.4)
Administrative expenses include the following:	
Employee costs	(231.1)
Occupancy costs	(77.5)
Repairs & maint. of property	(19.1)
Depreciation	(38.2)
Amortization	(2.7)
Other costs	(201.5)
	$ (570.1)

Current Guidance

Current guidance on reporting materiality requires consideration of quantitative and qualitative factors that determine whether it is "probable that the judgment of a reasonable person relying upon the report would have been changed" (SAB 99, SEC 1999). The fact that disaggregation by nature requires that management choose categories that will "enhance the usefulness of information in predicting the entity's cash flows" (Preliminary Views, para. 3.42) suggests that disaggregated amounts will affect the judgment of a reasonable investor, which is the overall basis for judging reporting materiality. If disaggregation results in greater use by investors and/or greater regulatory attention to individual line items that contain errors, then auditors should be more likely to require correction of a discovered error.

Concerning specific quantitative factors, current auditing standards are unclear on the effects of expense disaggregation. On the one hand, nearly all current audit guidance on reporting materiality includes the phrase "in relation to the statements taken as a whole." This could imply that statement totals and subtotals are the most important benchmarks. On the other hand, most reporting

materiality guidance also makes reference to one or more of the terms "account" or "disclosure," which suggests that disaggregated amounts are also important. The clearest reporting materiality guidance is provided in SAB 99, which requires that quantitative assessments be made "in relation to *individual line item amounts*, subtotals, or totals in the financial statements" (SAB 99, quoting AU 312.34; emphasis added).[5] PCAOB Auditing Standard No. 14 is less clear on this issue, requiring assessments to be made "in relation to the specific accounts and disclosures involved and to the financial statements as a whole" (PCAOB 2010a).[6] ISA 450 (IAASB 2010a) uses similar language ("in relation to particular classes of transactions, account balances or disclosures and the financial statements as a whole"). Both could be interpreted as supporting the importance of line items as a benchmark.[7] When existing audit guidance discusses *specific* quantitative materiality benchmarks related directly to the income statement, they tend to refer to subtotals such as gross profit, total expenses, or profit before tax (see e.g., ISA 320, IAASB 2010a). This again implies the primacy of subtotals and totals as benchmarks.

When there are multiple benchmarks, it is also not clear from the standards how the results from multiple benchmarks should be combined. The psychology literature has long recognized that such judgments might be made using combination rules ranging in complexity from one based solely on the benchmark most strongly indicating materiality to an importance-weighted combination of all benchmarks (e.g., Einhorn 1970).[8] Alternatively, a screening model based on the most important criteria might be used first, followed by a consideration of secondary criteria for borderline cases (e.g., Payne 1976). The combination rule employed can have a major impact on the resulting reporting materiality decision. Our general conclusion is that existing guidance, while less than perfectly clear, suggests that subtotals and totals are primary quantitative materiality benchmarks and, in some circumstances, *particularly important* accounts or line items are important but secondary benchmarks.[9] The effect of broad public income statement disaggregation of all (more and less important) expense items on materiality is not clear, particularly given that, as in our experiment, all auditors have detailed trial balance data available based on which materiality can be assessed.

Research Findings

Messier et al. (2005), Nelson (2004), Nelson and Tan (2005), and Libby and Seybert (2009) provide the most relevant recent reviews of literature examining quantitative and qualitative benchmarks used in practice for judging reporting materiality. Most recent studies show consistent focus on net income before tax as the most important quantitative benchmark, with materiality being defined as 5 percent of current period net income before tax.[10] As noted earlier, a \$0.02 per share error in Marks & Spencer's 2010 income statement or a \$0.17 per share error in Wal-Mart's

[5] This language is continued in the recently issued SAB 114 (SEC 2011b).

[6] PCAOB Auditing Standard No. 11, paragraph 7 includes some guidance related to planning materiality (PCAOB 2010a). This guidance requires auditors to assess the materiality of important accounts separately during the planning phase of the audit, and consider if lower materiality thresholds are appropriate. However, this guidance only applies to accounts that are especially important for the reporting company. Thus, we believe that this guidance relates to the type of broad disaggregation we examine in this study only if a disaggregated amount is judged to be particularly important to users.

[7] SAB 99 (SEC 1999) requires that auditors consider the effects on an operating segment as a qualitative factor.

[8] The first type of model is often called a conjunctive or multiple-cutoff model and the second a compensatory or trade-off model. Only in the second type of model can high materiality based on one benchmark be offset by low materiality on another (Libby 1981, 44–46).

[9] Our discussions with executive office representatives from the audit firms that participated in our experiment confirm this conclusion.

[10] In unusual periods, such as those where net income before taxes is very small or negative, other benchmarks such as sales and assets are used.

2010 income statement would not be quantitatively material using this benchmark. Following more recent regulation and guidance (SAB 99), errors below the quantitative criterion are subjected to an "SAB 99 review" for qualitative factors (e.g., Ng and Tan 2003). Recent studies of the transition to SAB 108 (SEC 2006), which requires that all discovered errors be evaluated based on both the "current period" ("rollover") and "cumulative" ("iron curtain") basis, strongly suggest that a large number of discovered quantitatively immaterial errors remain uncorrected (Keune and Johnstone 2009, 2012).[11] In addition, Joe et al. (2011) examine the period directly before the passage of the Sarbanes-Oxley Act and find that 24.2 percent of proposed adjustments are waived. Recent research by Acito et al. (2009) provides evidence that both qualitative and quantitative factors also influence how prior period errors are corrected.

Current research has little to say directly about the effects of broad expense disaggregation on the materiality of discovered misstatements. However, some research has examined how managers make disaggregation decisions. Heitzman et al. (2010) examine firms that disaggregate advertising expenses, finding that disaggregation is more likely when the information is more relevant to investors. Riedl and Srinivasan (2010) find that managers separately disclose special items when they are particularly relevant to investors. These findings support the idea that disaggregated line items affect users and could be considered when assessing materiality.

Findings from the voluntary disclosure and required segment disclosure literatures are indirectly related to the relevance of disaggregated amounts to users. Two studies suggest that voluntary disaggregation enhances the relevance of voluntary disclosures. Hirst et al. (2007) report that users believe that disaggregated management forecasts limit managers' ability to manage earnings and, as a result, find the forecasts to be more relevant. Supporting this idea, D'Souza et al. (2010) find that managers who regularly manage earnings limit disclosures of disaggregated information in the earnings press release. If auditors believe that the disaggregated amounts on the income statement will be used to assess management credibility, then they should serve as relevant materiality benchmarks. A number of studies examine managers' segment disaggregation decisions (e.g., Berger and Hann 2007; Botosan and Stanford 2005; Cho 2011) and demonstrate the effects of the disclosures on user decisions. To the degree that activities that respond similarly to similar economic forces are grouped together, segment reporting aids prediction of future cash flows (e.g., Berger and Hann 2003). Given that other forms of income statement disaggregation appear to have the potential to affect users judgments, to the degree that auditors believe that broad income statement disaggregation is aimed at improving users' ability to predict the entity's future cash flows, they could assume that disaggregated line items are also important benchmarks for judging materiality.

In addition, two studies address factors that could moderate the effect of expense disaggregation. Libby et al. (2006) find that audit partners require greater correction of misstatements in amounts recognized on financial statements than amounts disclosed in the notes because they view equivalent misstatements in recognized amounts as more material. Frederickson et al. (2006) note that, in their study of users, decision makers react to differences in secondary factors only if they are presented in a particularly transparent fashion. This could also be the case with U.S. auditors' consideration of line item amounts as benchmarks. This reasoning suggests that

[11] Keune and Johnstone (2009) report that 367 companies report uncorrected detected errors that were immaterial under either the income statement or balance sheet approaches as applied in prior years that became material under the new dual approach required by SAB 108. Although there is no way to directly measure the number of uncorrected immaterial errors that were still immaterial under the dual approach, we suggest that it is likely to be even larger given that the materiality measures under the income statement and balance sheet approaches are likely to be correlated. See Nelson et al. (2002) and Nelson et al. (2005) for evidence from before the issuance of SAB 108.

any effect of disaggregation will be greater if the amounts are presented in the statements as opposed to the notes.

While we expect note presentation to decrease the effect of disaggregation, there are two reasons why earlier results might not generalize to the current setting. First, Libby et al. (2006) ask auditors to judge the materiality of an error in numbers that are *either* recognized in the financial statements or disclosed in the notes. In the case of expense disaggregation, the amounts involved are always recognized on the income statement, and the only question is whether the underlying detail is presented on the face of the statements or the notes. The fact that the amounts *are* recognized could very well increase auditors' perceptions of their responsibilities for the more detailed amounts regardless of disclosure location. Also, auditors' responsibilities for the accuracy and completeness of note disclosures have received recent attention from audit standard-setters and regulators (e.g., IAASB 2010b; PCAOB 2011) that was not evident during the time period in which Libby et al. (2006) was conducted. This attention could also increase auditors' perceptions of their responsibilities for the note disclosures.

Hypotheses

The above discussion suggests that auditors' materiality decisions will be affected by the disaggregation of items on the income statement. However, the existence and magnitude of the effect is likely to depend on whether the disaggregated amounts are presented on the face of the statements or in the notes, and whether auditors naturally consider disaggregated amounts as primary benchmarks.

The experiment addresses the following specific hypotheses:

H1: Auditors will require correction of smaller errors when disaggregated expense item amounts are added to the face of the income statement in a between-subjects setting.

H2: Auditors will require correction of smaller errors when disaggregated expense item amounts are added to the face of the income statement as opposed to the notes in a between-subjects setting.

H3: Auditors will judge errors to be more material when disaggregated expense item amounts are added to the face of the income statement in a within-subjects setting.

H4: Auditors will judge errors to be more material when disaggregated expense item amounts are added to the face of the income statement as opposed to the notes in a within-subjects setting.

It also addresses the following research question:

RQ1: What considerations do auditors believe determine the effect of disaggregation of expense items on the materiality of discovered financial statement errors?

Our experiment tests the effects of disaggregation on both the face of the statements (H1 and H3) and in the notes (H2 and H4). It also tests these effects both in a between-subjects (H1 and H2) and a within-subjects setting (H3 and H4). As Libby et al. (2002, 804) note, "the between-subjects design provides a clean test of the subject's natural reasoning process, while the within-subjects design draws attention to the independent variable of interest and thus gives the subject a chance to detect and correct errors and inconsistencies in their responses." The between-subjects test is a better indicator of what participants do, while the within-subjects test is a better indicator of what they believe or know. Our experiment addresses these questions in a setting in which auditors rely exclusively on their knowledge of *current* reporting and auditing standards, without adding any regulatory requirements for income statement disaggregation.

III. EXPERIMENTAL PROCEDURES AND DESIGN

Participants

A total of 78 experienced U.S. auditors from three Big 4 firms participate in the study. Of the 78, 76 identify their current position as audit manager; two identify as audit seniors. Of the 78 participants, 37 complete the experiment online and 41 complete a pencil-and-paper version of the task.[12] The three firms provide between 14 and 41 auditors each. The auditors are selected and contacted directly by a senior member from the executive office of their firm. Most of the participants' experience is with public, commercial (nonfinancial), for-profit companies.[13] Ninety-seven percent (76 of 78) of participants have more than 6 years of total audit experience. The participants are assigned randomly to the experimental conditions.

Task

The participating auditors determine the materiality of a single audit difference involving the underaccrual of occupancy expenses. After reviewing an informed consent form, participants read background information about M&S, a publicly listed retailer with a moderate-sized analyst following. Participants learn that earnings have been growing steadily over the last five years, except for a dip in 2008, and that maintaining and increasing share price is an important management goal. To ensure that participants in all treatment conditions have access to the same account-by-account information, all participants receive a trial balance presenting the accounts in exactly the same level of disaggregation as presented in the disaggregated condition income statement, and have easy access to the trial balance during the between-subjects portion of the task. After reviewing the background information about the firm, participants are informed that one potentially important audit difference has been uncovered by the audit staff late in the audit. They are told that M&S leases the majority of its store space through operating leases, and that M&S miscalculated occupancy costs related to selling and marketing expenses, resulting in a $79.3 million underaccrual that overstates net income by $79.3 million ($51.5 million) before (after) tax. The error is approximately 4.5 percent of pretax income, 3.8 percent of selling and marketing expenses (disclosed to the public in all conditions), and 18.1 percent of occupancy costs (disclosed to the public only in the disaggregated conditions).[14] Based on consultation with executive office partners from the participating firms, we chose this error amount because it is close to, but below, the most commonly used quantitative materiality benchmark of 5 percent of net income before tax.[15] It also is a larger percentage of occupancy costs than the relevant subtotals (selling and marketing expenses) and totals (net income). As such, the error provides the greatest opportunity for additional disaggregation to influence materiality judgments.

Participants are further informed that the error is inadvertent and no other qualitative materiality factors are present, and that the client objects to any adjustments of the statements and disclosures. Without the latter statement, auditors would perceive no need to consider waiving adjustment. We inform participants that no other qualitative factors are present to help ensure that

---

[12] Administration and firm do not affect the main results.

[13] Seventy-two percent of their experience is with public companies, 26 percent with private companies, and 2 percent with not-for-profit organizations. Seventy-eight percent (61 of 78) have no experience in auditing financial institutions and 79 percent (62 of 78) have 20 percent or less of their public company audit experience with retailers (responses to these questions are categorical).

[14] These percentages are not calculated for the participants to allow them to choose their own materiality benchmarks.

[15] Executive office partners from each of the participating firms provided detailed comments on the experimental instrument. Their recommendations were consistent across firms.

FIGURE 2
Experimental Manipulations

Panel A: Information Presented to Participating Auditors in the Aggregated Condition

Required: Your task is to determine what the final outcome of an audit would be with respect to the audit difference. To help you consider these alternatives, the tables below show ***how the income statement would appear to users in the Form 10-K if "no adjustment" and "full adjustment" is made***, with the affected areas highlighted.

Audit Difference: Occupancy costs related to selling and marketing are understated by $79.3 million before tax ($51.5 million after tax); all indications are that the error is inadvertent M&S does not separately report occupancy costs on the face of the income statements or notes.

If NO ($0) adjustment is made, the audited statement would appear as follows (in millions):	
Net sales	$ 12,262.1
Cost of sales	(7,690.2)
Gross profit	4,571.9
Selling and marketing expenses	(2,074.4)
Administrative expenses	(570.1)
Operating profit	1,927.4
Interest, net	(164.5)
Income before income taxes	1,762.9
Income tax expense	(617.0)
Net income	$ 1,145.9

If FULL ($79.3) adjustment is made, the audited statement would appear as follows (in millions):	
Net sales	$ 12,262.1
Cost of sales	(7,690.2)
Gross profit	4,571.9
Selling and marketing expenses	(2,153.7)
Administrative expenses	(570.1)
Operating profit	1,848.1
Interest, net	(164.5)
Income before income taxes	1,683.6
Income tax expense	(589.3)
Net income	$ 1,094.3

(continued on next page)

the level of aggregation is the primary driver of our results. Participants are then told about the presentation of occupancy cost in the client's financial statements, which is the independent variable in our design, and are asked a manipulation check question to ensure their attention to this matter. Participants are told that this is a known misstatement in the financial statements.[16] AU 312 (AICPA 2010) requires auditors to notify management of known misstatements and request that management correct them. However, the standard does not require auditors to *force* their clients to correct them.[17]

The following page repeats the specific facts about the error, and provides participants with the "full correction" and "no correction" financial statements and notes (where applicable) shown in Figure 2, Panels A, B, or C (depending on experimental condition).[18] This presentation helps ensure that participants understand how correcting or not correcting the error would affect the company's disclosures in their experimental condition. Beneath this figure, the participants' materiality

[16] We conjecture that our participants would have judged a "likely" misstatement to have higher allowable error, but the effect of disaggregation on the judgment would have been similar.

[17] AU 312 no longer officially applies to U.S. public company audits. Neither PCAOB Auditing Standard No. 11 nor 14 reference known errors. At a higher level, auditors would have to consider whether any failure to correct a known error was evidence of intentional manipulation under SAB 99.

[18] The expense data were identical to that provided in the 2009 Marks & Spencer financial report except that occupancy costs related to Administrative expenses were merged with other costs to avoid any confusion about the location of the error.

FIGURE 2 (continued)

Panel B: Information Presented to Participating Auditors in the Disaggregated-Face Condition

Required: Your task is to determine what the final outcome of an audit would be with respect to the audit difference. To help you consider these alternatives, the tables below show ***how the income statement would appear to users in the Form 10-K if "no adjustment" and "full adjustment" is made***, with the affected areas highlighted.

Audit Difference: Occupancy costs related to selling and marketing are understated by $79.3 million before tax ($51.5 million after tax); all indications are that the error is inadvertent. M&S separately reports occupancy costs on the face of the income statement.

If NO ($0) adjustment is made, the audited statement would appear as follows (in millions):	
Net sales	$ 12,262.1
Cost of sales	(7,690.2)
Gross profit	4,571.9
Selling and marketing expenses	
Employee costs	(923.2)
Occupancy costs	(439.2)
Repairs & maint. of property	(76.6)
Depreciation	(343.5)
Amortization	(24.6)
Other costs	(267.3)
	(2,074.4)
Administrative expenses	
Employee costs	(231.1)
Repairs & maint. of property	(19.1)
Depreciation	(38.2)
Amortization	(2.7)
Other costs	(279.0)
	(570.1)
Operating profit	1,927.4
Interest, net	(164.5)
Income before income taxes	1,762.9
Income tax expense	(617.0)
Net income	$ 1,145.9

If FULL ($79.3) adjustment is made, the audited statement would appear as follows (in millions):	
Net sales	$ 12,262.1
Cost of sales	(7,690.2)
Gross profit	4,571.9
Selling and marketing expenses	
Employee costs	(923.2)
Occupancy costs	(518.5)
Repairs & maint. of property	(76.6)
Depreciation	(343.5)
Amortization	(24.6)
Other costs	(267.3)
	(2,153.7)
Administrative expenses	
Employee costs	(231.1)
Repairs & maint. of property	(19.1)
Depreciation	(38.2)
Amortization	(2.7)
Other costs	(279.0)
	(570.1)
Operating profit	1,848.1
Interest, net	(164.5)
Income before income taxes	1,683.6
Income tax expense	(589.3)
Net income	$ 1,094.3

(continued on next page)

judgments are captured in two questions. They first answer "Would the failure to correct this error make the financial statements materially misstated?" on a six-point Likert-type scale, with the endpoints labeled "Definitely NOT materially misstated" and "Definitely MATERIALLY MISSTATED" (the Likert scale measure). Participants then answer our primary measure: "What is the maximum dollar understatement of pretax occupancy costs that could exist without making the financial statements materially misstated?" (the allowable error measure). This measure is our primary measure because it has a real empirical referent and is a more precise scale.

Participants then assess the extent to which auditors would pressure the client to correct this error, and are asked a manipulation check question about the format of the financial statements in the case and a comprehension check question about the size of the selling and marketing expenses account. Participants then respond to a series of within-subjects questions (described in the next

FIGURE 2 (continued)

Panel C: Information Presented to Participating Auditors in the Disaggregated-Notes Condition

Required: Your task is to determine what the final outcome of an audit would be with respect to the audit difference. To help you consider these alternatives, the tables below show ***how the income statement would appear to users in the Form 10-K if "no adjustment" and "full adjustment" is made***, with the affected areas highlighted.

Audit Difference: Occupancy costs related to selling and marketing are understated by $79.3 million before tax ($51.5 million after tax); all indications are that the error is inadvertent. M&S separately reports occupancy costs in the notes, but not on the face of the income statement

If NO ($0) adjustment is made, the audited statement would appear as follows (in millions):	
Net sales	$ 12,262.1
Cost of sales	(7,690.2)
Gross profit	4,571.9
Selling and marketing expenses	(2,074.4)
Administrative expenses	(570.1)
Operating profit	1,927.4
Interest, net	(164.5)
Income before income taxes	1,762.9
Income tax expense	(617.0)
Net income	$ 1,145.9
Notes to Financial Statements:	
Selling and marketing expenses include the following:	
Employee costs	(923.2)
Occupancy costs	(439.2)
Repairs & maint. of property	(76.6)
Depreciation	(343.5)
Amortization	(24.6)
Other costs	(267.3)
	$ (2,074.4)
Administrative expenses include the following:	
Employee costs	(231.1)
Repairs & maint. of property	(19.1)
Depreciation	(38.2)
Amortization	(2.7)
Other costs	(279.0)
	$ (570.1)

If FULL ($79.3) adjustment is made, the audited statement would appear as follows (in millions):	
Net sales	$ 12,262.1
Cost of sales	(7,690.2)
Gross profit	4,571.9
Selling and marketing expenses	(2,153.7)
Administrative expenses	(570.1)
Operating profit	1,848.1
Interest, net	(164.5)
Income before income taxes	1,683.6
Income tax expense	(589.3)
Net income	$ 1,094.3
Notes to Financial Statements:	
Selling and marketing expenses include the following:	
Employee costs	(923.2)
Occupancy costs	(518.5)
Repairs & maint. of property	(76.6)
Depreciation	(343.5)
Amortization	(24.6)
Other costs	(267.3)
	$ (2,153.7)
Administrative expenses include the following:	
Employee costs	(231.1)
Repairs & maint. of property	(19.1)
Depreciation	(38.2)
Amortization	(2.7)
Other costs	(279.0)
	$ (570.1)

section) and answer a series of demographic questions. The web-based version of the experiment prohibits referral to earlier portions of the experiment. The paper-and-pencil version enforces the same restrictions through use of three separate sealed envelopes.[19]

Between-Subjects Design

The experiment uses a 1 × 3 between-subjects design, varying the level of disaggregation included in the client's financial statements.

[19] In the pencil-and-paper version of the experiment, the debriefing questions and within-subjects questions are answered after sealing the between-subjects task in envelope 1. These questions are then sealed in envelope 2 before the demographic questions are answered and placed back in envelope 3.

In the aggregated condition (shown in Figure 2, Panel A), participants view an income statement that includes the typical level of disaggregation that is included on U.S. retailer financial statements (net sales, cost of sales, gross profit, selling and marketing expenses, administrative expenses, operating profit, interest, income before income taxes, income tax expense, and net income). In the disaggregated-face condition (shown in Figure 2, Panel B), the same financial statement line items are included, but the income statement includes details about the six individual expense amounts that make up the selling and marketing expenses subtotal and the five that make up the administrative expenses subtotal, with individual expense amounts disaggregated by nature. In the disaggregated-notes condition (shown in Figure 2, Panel C), participants view an income statement that is identical to the aggregated condition, and are also shown a note to the financial statements that includes the details described above. The totals and subtotals included in each set of financial statements are identical across conditions.

The discovered audit difference included in the experimental materials is in one of the line items listed in the disaggregated conditions as well as the trial balance shown to all participants. As a consequence, the total information available to the participants is identical across conditions. Participants are always free to refer back to the trial balance during the between-subjects portion of the experiment. As a consequence, participants in all conditions are aware of the nature and amount of the trial balance account that is affected, which should, according to standards, serve as a benchmark for audit planning. Our experimental manipulation only changes the level of disaggregation or the location of disaggregated information that is included in the client's public financial statements.

Within-Subjects Design

In the within-subjects portion of the experiment, participants are told to "Assume that Companies A and B present the following income statements (and related notes) in their Form 10-K. Further assume that both companies miscalculate occupancy costs related to selling and marketing, which results in an underaccrual of $79.3 million. The only difference between the companies' income statements is that Company B disaggregates the elements of selling and marketing expenses and administrative expenses. The companies are equal in all other respects." Participants in the two disaggregated conditions compare the statements (and notes in the disaggregated-notes condition) shown in their between-subjects task to the aggregated statements. Participants in the aggregated condition compare the aggregated condition statements to those presented in the disaggregated-face condition. Company A always presents the aggregated statements.

Figure 3 shows the statements presented in the aggregated and disaggregated-face conditions. As noted above, differences in these within-subjects comparisons indicate participants' beliefs about the implications of the differences in aggregation, whereas the between-subjects comparisons better reflect their natural reasoning process.

The participating auditors compare the two financial statement presentations on four, nine-point Likert-type scales with 1 labeled "More likely effect for Company A," 5 labeled "No difference in likelihood of effect," and 9 labeled "More likely effect for Company B," with small changes to the wording of the labels to reflect the question:

1. For which company would the error in occupancy costs be more material? (the materiality question)
2. For which company would a reasonable investor's judgments be more likely to be affected by a failure to correct the error in occupancy costs? (the investor judgment question)

FIGURE 3

Financial Statements Presented to Participants in the Within-Subjects Portion

For the following 4 questions, rely on the same error and same data for Companies A and B.

Assume that Companies A and B present the following income statements in their Form 10-K. Further assume that both companies miscalculated occupancy costs related to selling and marketing which resulted in an underaccrual of $79.3 million. The only difference between the companies' income statements is that Company B disaggregates the elements of selling and marketing expenses and administrative expenses. The companies are equal in all other respects.

Company A	
Net sales	$ 12,262.1
Cost of sales	(7,690.2)
Gross profit	4,571.9
Selling and marketing expenses	(2,074.4)
Administrative expenses	(570.1)
Operating profit	1,927.4
Interest, net	(164.5)
Income before income taxes	1,762.9
Income tax expense	(617.0)
Net income	$ 1,145.9

Company B	
Net sales	$ 12,262.1
Cost of sales	(7,690.2)
Gross profit	4,571.9
Selling and marketing expenses	
Employee costs	(923.2)
Occupancy costs	(430.2)
Repairs & maint. of property	(76.6)
Depreciation	(343.5)
Amortization	(24.6)
Other costs	(267.3)
	(2,074.4)
Administrative expenses	
Employee costs	(231.1)
Repairs & maint. of property	(19.1)
Depreciation	(38.2)
Amortization	(2.7)
Other costs	(279.0)
	(570.1)
Operating profit	1,927.4
Interest, net	(164.5)
Income before income taxes	1,762.9
Income tax expense	(617.0)
Net income	$ 1,145.9

3. For which company would the SEC be more likely to require a restatement to correct the error in occupancy costs? (the SEC attention question)
4. For which company's statements would it take more hours for the auditors to complete their audit work? (the audit hours question)

Each question also asked for a free-response explanation to the question, "Why?" The participants are also asked for any additional comments on audit effects of disaggregation.

IV. RESULTS

Between-Subjects Results

Manipulation Check

In debriefing, 97 percent (76 of 78) of participants correctly identify where occupancy costs are disclosed. The occupancy cost account comprehension check question requires participants to select an answer from six different categories. Twelve percent (9 of 78) of the participants indicate the amount of the error as their answer and 5 percent (4 of 78) of participants provide an answer that is more than one category away from the correct amount. There are no significant differences in

responses between conditions. Our inferences are unchanged if we exclude individuals who fail either check.[20]

Tests of H1 and H2

We test H1 and H2 with planned comparisons. All t-tests are one-tailed. Table 1, Panel A includes descriptive statistics. None of the differences on the Likert scale materiality question are significant (see Table 1, Panel B). As noted earlier, we expect the allowable error metric to provide a more precise measure of the dependent variable of interest because it has an empirical referent and is a finer scale. The remaining discussion of H1 and H2 relies on the allowable error judgments.[21]

We exclude two individuals who provided responses on the allowable error question that are more than three standard deviations above the mean. We also exclude one individual who responded that the appropriate allowable error is below 1.[22] The equality of variance assumption is violated for our allowable error data (Levene's $W = 8.56$, $p < 0.001$). Thus, we perform separate variance Bonferroni adjusted t-tests to evaluate our hypotheses reported in Table 1, Panel C.[23]

The mean allowable error in the disaggregated-face condition (mean = 68.93) is significantly lower ($t = 2.22$, $p = 0.032$) than in the aggregated condition (mean = 82.79), supporting H1. This result suggests that disaggregated expense item amounts on the face of the income statement increase auditors' judgments of the materiality of discovered financial statement errors.

Allowable error in the disaggregated-face condition (mean = 68.93) is lower than in the disaggregated-notes condition (mean = 81.05) by a marginally significant amount ($t = 1.78$, $p = 0.082$). This test supports H2 indicating that disaggregating expense item amounts on the face of the statements as opposed to the notes increases auditors' judgments of the materiality of discovered financial statement errors[24]

---

[20] We also examine the amount of time auditors take to complete the experiment. We collected timing data for the 37 auditors who completed the experiment online. These auditors took an average of 24.7 minutes to complete the online instrument, and spent 9.4 minutes on the between-subjects portion of the instrument. There are no differences in time spent between experimental conditions.

[21] We adjust for unequal variances in our parametric tests. The descriptive statistics suggest that the normality assumption is also violated for the allowable error measure. However, our relatively large sample size ($n = 75$ for the three treatments) suggests that parametric tests that adjust for unequal variances should be robust to this deviation from the normality assumption (Scheffé 1959). Therefore, our analyses focus on means and we perform parametric tests. Using less-powerful nonparametric tests, the difference between the aggregated condition and disaggregated-face condition is significant at $p = 0.07$ Van der Waerden test and at $p = 0.11$ Mann-Whitney U test (Conover 1999).

[22] The individual who reported an "allowable error" amount under 1 is indicating that there is no materiality threshold—that all errors must be corrected. This is not consistent with the materiality guidance and would produce an unreasonably inefficient audit. The two individuals who reported "allowable error" more than 3 standard deviations above the mean are indicating that their materiality threshold is more than twice 5 percent of net income, a materiality threshold that would be unacceptable to a reviewing partner or other authority within the auditing firm. Thus, we believe that these responses should be excluded from our analysis. However, when we include these three extreme individuals in our analysis, we treat the responses as a binary variable—either greater than or equal to versus less than the amount of the discovered error. The difference between the disaggregated-face and aggregated conditions is marginally significant ($p = 0.094$, one-tail) using the Fisher exact probability test. No other differences are significant. This test provides additional comfort that our main result is robust to inclusion of these extreme observations.

[23] We use the Bonferroni adjustment of 2 to our significance values because we conduct two paired comparisons using the aggregated group data.

[24] The Brown Forsythe unequal variances ANOVA across all three conditions is significant at $p = 0.0513$. Allowable error in the disaggregated-notes condition is not significantly higher than in the aggregated condition ($t = 0.34$, $p = 0.73$) in an additional unequal variance t-test.

TABLE 1

Descriptive Statistics and Planned Comparisons for Between-Subjects Responses

Panel A: Descriptive Statistics—Mean, (Standard Deviation), and [Median]

	Experimental Condition			
	Aggregated	**Disaggregated-Face**	**Disaggregated-Notes**	**Overall**
Likert Scale Judgments	3.22	3.41	2.96	3.2
	(1.54)	(1.69)	(1.51)	(1.58)
	[3]	[3]	[2]	[3]
	n = 23	n = 27	n = 25	n = 75
Allowable Error	82.79	68.93	81.05	77.22
	(14.56)	(28.39)	(20.45)	(22.84)
	[88.00]	[84.20]	[88.00]	[87.00]

Panel B: Likert Scale Materiality Judgments: Bonferroni Adjusted Pooled Variance t-tests

Hypothesis	Comparison	Difference	df.	t-value	p-value
1	Disaggregated-Face > Aggregated	0.19	72	0.42	0.674
2	Disaggregated-Face > Disaggregated-Notes	0.45	72	1.02	0.313

Panel C: Allowable Error Judgments: Bonferroni Adjusted Separate Variance t-tests

Hypothesis	Comparison	Difference	df.	t-value	p-value
1	Disaggregated-Face < Aggregated	−13.86	40	2.22	0.032
2	Disaggregated-Face < Disaggregated-Notes	−12.12	47	1.78	0.082

This table presents descriptive statistics and Bonferroni Adjusted Separate Variance t-tests for the materiality Likert scale and allowable error measures included in our experiment. In this experiment, 75 auditors evaluate the materiality of an error on a six-point Likert-type scale, with the endpoints labeled "Definitely NOT materially misstated" (1) and "Definitely MATERIALLY MISSTATED" (6) (the Likert scale measure). Participants are also asked to provide the maximum dollar understatement of pretax occupancy costs that could exist without making the financial statements materially misstated (the allowable error measure). Lower allowable error implies higher materiality. The experiment manipulates the level of expense disaggregation included within the company's financial statements: aggregated presentation (no disaggregation), disaggregated presentation on the face of the income statement, and disaggregated presentation in the notes of the financial statements. t-tests are one-tailed.

Discussion of Results

Taken together, the results show that disaggregated information on the face of the income statement increases the materiality of errors affecting the disaggregated accounts. However, when disaggregated information is included in the notes, this effect is reduced. The significant difference in variances for allowable error also suggests that there is more disagreement (less consensus) on the appropriate materiality standard when disaggregated expense items are presented on the face of the statements.

Within-Subjects Results

This section evaluates the responses that the auditors provided for the four within-subjects questions included in our experimental instrument. Within-subjects questions make the difference

salient and more directly assess participants' beliefs about the appropriate effect of the inclusion of disaggregated expense amounts in financial statements. All auditors compare a company with aggregated expenses with an otherwise identical company with disaggregated expenses on nine-point Likert-type scales. On each scale, the response 5 is labeled as "No difference in likelihood of effect," and a response above 5 means a more likely effect for the company with disaggregated statements. Table 2, Panel A presents descriptive statistics by treatment.[25]

Tests of H3 and H4

To evaluate H3, we examine responses to the materiality within-subjects question provided by the 53 participants in the aggregated and disaggregated-face conditions (labeled "Combined" in Panel A) that compare aggregated statements to disaggregated-face statements. Table 2, Panel B shows that the mean response to the materiality question of 6.25 is significantly above 5 ($t = 6.69$, $p < 0.001$). This comparison indicates that auditors believe disaggregated expense item disclosure on the face of the statements would increase the materiality of the error and supports H3. None of the auditors believe that disaggregation would decrease the materiality of the error. Fifty-eight percent (31 of 53) of the participating auditors suggest that disaggregating expenses would increase the materiality of the error; 42 percent (22 of 53) suggest it will have no effect. While, on average, auditors believe that disaggregated expense items are relevant materiality benchmarks, there is little consensus on this issue.

H4 predicts that the effect of disaggregated expense item amounts in H3 will be greater if the amounts are placed on the face of the income statement as opposed to in the notes. We test this hypothesis by comparing the answers to the materiality within-subjects question between the 53 participants comparing the aggregated statements with disaggregation on the face of the statements (labeled "Combined") to the 25 participants comparing the aggregated statements to those with disaggregation in the notes. As shown in Table 2, Panel C, this difference in means is significant ($t = 2.22$, $p = 0.016$). Forty percent (10 of 25) of the auditors comparing aggregated statements to statements with disaggregation in the notes indicate that disaggregating expenses would increase the materiality of the error. As described above, 58 percent of auditors comparing aggregated statements to statements with disaggregation in the face indicate that disaggregating expenses would increase the materiality of the error. This difference in proportions is marginally significant (untabluated, Fisher exact probability, $p = 0.100$). Consistent with the results of the between-subjects tests, auditors believe that disaggregation in the notes reduces the relevance of disaggregated numbers as a materiality benchmark.

We also examine participating auditors' responses to the other three within-subjects questions. Table 2, Panel B shows that all the means for the comparisons between the aggregated and disaggregated-face statements are significantly above 5, providing evidence that auditors believe that investor judgments are more likely to be affected by the error, the SEC is more likely to require restatement, and audit hours are likely to increase when disaggregated expense item disclosures are included in financial statements.

We also compare responses to the other three within-subjects questions from the auditors comparing the aggregated statements to those with disaggregation on the face of the statements to those comparing the aggregated statements to those with disaggregation in the notes. Table 2, Panel C reports these results. The only significant difference is on the SEC attention question ($t = 3.12$, $p = 0.001$) with SEC attention increasing more with disaggregation on the face of the statements than with disaggregation in the notes.

---

[25] All participants are included in this table and the related analysis. If the three subjects excluded from the between-subjects tests are excluded, all statistical inferences are unchanged.

TABLE 2

Descriptive Statistics and Statistical Tests for Within-Subjects Responses

Panel A: Descriptive Statistics by Experimental Condition—Mean, (Standard Deviation), and [Number of Subjects Providing Answers above 5]

Between-Subjects Experimental Condition	**Aggregated**	**Disaggregated-Face**	**Combined**	**Disaggregated-Notes**	
Within-Subjects Comparison	**Aggregated versus Disagg.-Face**	**Aggregated versus Disagg.-Face**	**Aggregated versus Disagg.-Face**	**Aggregated versus Disagg.-Notes**	**Overall**
Materiality	6.54	6.00	6.25	5.64	6.06
	(1.49)	(1.22)	(1.36)	(1.03)	(1.29)
	[17]	[14]	[31]	[10]	[41]
	n = 25	n = 28	n = 53	n = 25	n = 78
Investor Judgments	6.22	5.78	5.99	6.16	6.04
	(1.31)	(1.20)	(1.26)	(1.21)	(1.24)
	[16]	[11]	[27]	[15]	[42]
	n = 25	n = 28	n = 53	n = 25	n = 78
SEC Attention	6.58	6.29	6.42	5.60	6.16
	(1.34)	(1.51)	(1.45)	(0.87)	(1.34)
	[18]	[14]	[32]	[10]	[42]
	n = 25	n = 28	n = 53	n = 25	n = 78
Audit Hours	6	5.61	5.79	6.24	5.94
	(1.22)	(0.92)	(1.08)	(1.42)	(1.21)
	[14]	[12]	[26]	[15]	[41]
	n = 25	n = 28	n = 53	n = 25	n = 78

Panel B: Test of Mean Responses to Within-Subject Questions > 5: Aggregated versus Disaggregated-Face

Dependent Variable	**Difference from 5**	**df.**	**t-value**	**p-value**
Materiality	1.25	52	6.69	< 0.0001
Investor Judgments	0.99	52	5.76	< 0.0001
SEC Attention	1.42	52	7.14	< 0.0001
Audit Hours	0.79	52	5.34	< 0.0001

Panel C: t-tests of Differences on Within-Subjects Responses by Experimental Condition

Variable	**Comparison**	**Difference**	**df.**	**t-value**	**p-value**
Materiality	Disaggregated-Face > Disaggregated-Notes	0.61	61	2.22	0.016
Inv Judgments	Disaggregated-Face > Disaggregated-Notes	−0.16	76	−0.56	ns
SEC Attention	Disaggregated-Face > Disaggregated-Notes	0.82	72	3.12	0.001
Audit Hours	Disaggregated-Face > Disaggregated-Notes	−0.45	76	−1.54	ns

(continued on next page)

TABLE 2 (continued)

This table presents descriptive statistics and statistical tests of the responses that participating auditors provided to our within-subjects questions. In this portion of the experiment, participants are told that an identical error exists in the financial records of two companies, Company A and Company B. Company A presents aggregated information in its income statement, while Company B presents disaggregated information (on either the face of the income statement or in the notes, depending on experimental condition); they are equal in all other respects. Participants evaluate the materiality of the error (materiality), the likelihood of the error influencing an investor's judgments (investor judgments), the likelihood that the SEC would require adjustment (SEC attention), and the amount of audit hours required to complete the audit (audit hours). All questions are answered on nine-point Likert-type scales with 1 labeled "More likely effect for Company A," 5 labeled "no difference in likelihood of effect," and 9 labeled "More likely effect for Company B," with small changes to the wording of the labels to reflect the question. t-tests are one-tailed.

Panel A presents descriptive statistics for each measure.

Panel B presents a t-test that evaluates if the mean responses provided to each of the four within-subjects questions are different from 5 in the combined aggregated and disaggregated-face conditions where the aggregated and disaggregated-face statements were compared.

Panel C presents t-tests comparing the mean responses on each within-subject question in the Aggregated and Disaggregated-Face conditions (combined) with the Disaggregated-Notes condition. We perform this comparison because individuals in the Aggregated and Disaggregated-Face condition reviewed the same information on a within-subjects basis. Note that equality of variance assumption is violated for the Materiality (Levine's $W = 3.62$, $p = 0.03$) and the SEC Attention (Levine's $W = 7.01$, $p = 0.002$) variables. Thus, this table presents separate variance t-tests for the materiality and SEC attention variables; investor judgments and audit hours are evaluated using a pooled variance t-test.

Test of RQ1

We also ask the open-ended question "Why?" after each within-subjects question. We analyze the answers to determine what considerations drive participating auditors' judgments when considering discovered financial statement errors in statements that include disaggregated expense item amounts. Since the implications of existing audit guidance related to the disaggregation in our case is somewhat unclear, we do not categorize the responses as correct or incorrect. Instead, we review the within-subjects responses and develop a coding scheme that groups explanations into broad categories. We then code responses according to this scheme and analyze the results. The categories and descriptive statistics for all four free-response questions are included in Table 3, Panels A and B.

We first examine "Why?" responses to the materiality within-subjects question. Of the 78 participants, 73 provide answers to the open-ended questions. Of the 73, 35 report that disaggregation would *not* influence their materiality assessments on the within-subjects portion of our experiment. The majority (18) of these 35 auditors mention that overall materiality is determined at an aggregated level (coded as 2). Typical responses include:

> "Materiality is based on income before income taxes."

> "Auditors do not opine on separate line items in financial statements ... the error is material to the financials taken as a whole."

> "Total selling and marketing costs are the same. Subtotals within that group have little relevance in determining materiality. Pre-tax income is the main driver of materiality."

Thirty-eight auditors report that disaggregation *would* influence materiality. Among these auditors, the most common informative response is a reaction to perceived investor needs (coded as 5), suggesting that auditors view voluntary disclosures in a manner similar to Heitzman et al.'s (2010) view. Typical responses include:

> "Increased disaggregation may lead to increased reader/investor reliance or use of the additional information."

TABLE 3

Coding Scheme and Descriptive Statistics for Written Responses

Panel A: Description of Coding Scheme

We reviewed each of the written responses, and used the following coding scheme to evaluate our results:

#1	Discussion of the irrelevance of line item disaggregation
#2	Discussion that materiality is determined at an aggregated level
#3	Discussion that materiality is determined by size/impact of the error
#4	Discussion of disaggregation/line item error without any additional information
#5	Discussion of the importance of the line item to users of financial statements
#6	Discussion of displaying line items on the face of the I/S
#7	Discussion of visibility of the line item to users of financial statements
#8	Discussion of evaluation of the error as a portion of a specific line item
#9	Discussion of auditing at a disaggregated level regardless of presentation
#10	Additional hours for auditing classification/components
#11	Additional hours for tying-out financial statements
#12	Additional hours generated by having more line items
#13	Other rationale

Classifications 1 through 3 capture the typical responses provided by individuals who report that disaggregation will not influence their materiality judgments. These answers are commonly provided for every within-subjects question.
Classifications 4 through 9 capture the reasons participants provide for claiming that disaggregation matters. Classifications 4 through 7 are used to classify responses provided for the Materiality, Investor Judgments, and SEC Attention questions, respectively. Classification 8 is included to capture answers to the Investor Judgments question, while Classification 9 is included to capture answers to the SEC Attention question.
Classifications 10 through 13 are used to code answers to the Hours variable. Note that the answers provided to this question are often significantly different from answers provided to the other three questions.

Panel B: Descriptive Statistics for Written Responses

	Within-Subjects Variable				
	Materiality	Investor Judgments	SEC Attention	Audit Hours	Overall
Response # 1	10	16	7	1	34
Response # 2	18	5	9	5	37
Response # 3	7	11	8	4	30
Response # 4	18	10	10	1	39
Response # 5	12	19	6	1	38
Response # 6	4	2	3	0	9
Response # 7	1	8	1	0	10
Response # 8	0	0	15	0	15
Response # 9	0	0	0	8	8
Response # 10	0	0	0	8	8
Response # 11	0	0	0	10	10
Response # 12	0	0	0	18	18
Response # 13	3	2	8	9	22
Total	73	73	67	65	278

"Company B elected to present occupancy costs as a separate line item in the I/S, so it presumably carries some weight with users from the perspective of Company B's management. Given the perceived importance of this line item, the auditor would be more sensitive to a misstatement of that individual line item."

> "The fact that B disaggregated the occupancy costs suggests that the line item may be important to investors. More facts would be needed such as whether this line item is discussed in earnings releases conference calls."

These responses again suggest little consensus about the relevance of line items as materiality benchmarks. The following two quotes even indicate different interpretations of the same, admittedly confusing, professional guidance.

> "SAB 99 does not require management or the auditor to specifically evaluate an error against a particular financial statement line item. Therefore, I do not believe there is a difference in the 2 examples."[26]

> "Materiality to the line item should be considered although I don't believe it should be the only factor to be considered (just one of)."

The responses to these questions directly relate to the results of our tests of H1 and H2. Auditors' allowable error judgments depend critically on their interpretation of general (effects or lack of effect on users) or specific materiality guidance (importance of accounts or line items versus the statements taken as a whole). Furthermore, there is substantial disagreement on these issues. Responses to the open-ended questions on user and SEC attention are similar.

Additional Analysis

The analysis presented above suggests that the findings are a result of conscious differences in beliefs about the relevance of line items as materiality benchmarks. To provide greater insight into this issue, we partition our participants who compared the aggregated statements with disaggregation on the face of the statements in the within-participants task between those who suggested that disaggregating expenses on the face of the statement *would* increase materiality and those who suggested it *would not*, and examine differences in their between-subjects responses. This untabulated analysis indicates no between-participants difference on the Likert-type materiality scale between the aggregated or disaggregated-face conditions for either the "would" ($t = 0.77$, $p = 0.224$) or the "would not" ($t = -0.42$, $p = \text{ns}$) affect materiality group. When we perform the same analysis using the allowable error measure, the effects of aggregation are significant for the "would" affect materiality group ($t = 2.11$, $p = 0.024$) and nonsignificant ($t = 0.87$, $p = 0.196$) for the "would not" affect materiality group.[27] This suggests that the lack of findings for our tests of H1 using the between-subjects materiality Likert-type scale is most likely the result of power issues. It also adds further support for the conclusion that our allowable error results are the result of conscious differences in beliefs.

Discussion of Results

The within-subjects analysis and open-ended questions provide a strong indication of what participants believe about the task at hand. These tests show that, on average, including disaggregated line item amounts in financial statements increases the materiality of discovered financial statement errors in a within-subjects setting. In addition, our tests show that expectations that the error will influence investor judgments, expectations that the SEC would require

---

[26] Note that SAB 99 does require evaluation of an error against a particular financial statement line item.

[27] Tests are pooled variance for the Likert-type materiality scale and separate variance for the allowable error measure. Note that this biases the results against our findings. All p-values are one-tailed. These tests include all participants from the aggregated and disaggregated-face conditions included in the main between-participants tests.

restatement to correct the error, and expected audit hours are also affected in a within-subjects setting. We also find that the effect of disaggregation on responses to the materiality and SEC attention questions are stronger when the disaggregation occurs on the face of the financial statements. However, there is little agreement on the appropriate response or the interpretation of existing guidance.

V. CONCLUSION

Prior audit materiality research summarized by Messier et al. (2005) and others documents a growing consensus among auditors as to the appropriate quantitative and qualitative benchmarks used to evaluate the materiality of discovered financial reporting errors. These studies, our own results, and our discussions with representatives from major firms suggest the prominence of "5 percent of pre-tax income" as the key quantitative benchmark in normal times. Smaller errors over a certain amount are subjected to what many call a "SAB 99 review" for qualitative factors. The auditor error correction studies discussed earlier recognize that auditor error correction requires that the auditor detect the error and that management agrees to correct the error (or that the auditor forces the issue when management objects). In the current regulatory environment, many discovered errors are corrected voluntarily. But when management objects to correction, materiality considerations become paramount and recent research suggests that many discovered immaterial errors are not corrected. Our study adds to the literature that examines how the nature of financial reporting standards and financial statement format and disclosure location can affect auditors' materiality judgments. These materiality judgments ultimately affect the reliability of the reported numbers.

Our study suggests that disaggregating expense items can reduce the allowable error in the disaggregated amounts, increasing the reliability of the disaggregated amounts as well as the resulting statement subtotals and totals. However, there is significant disagreement among practicing auditors on the relevance of line items as materiality benchmarks and that reporting disaggregated amounts in the notes reduces the effect. This suggests that voluntary disaggregation decreases the average amount of error tolerated in current financial statements, but at the same time decreases the consensus in audit practice. The effect is substantially reduced if the disaggregated data are presented in the notes. Our within-subjects results suggest that this results from conscious differences in beliefs about the relevance of line items as materiality benchmarks. Both of these findings are relevant to financial reporting standard-setters and regulators interested in the effects of financial statement presentation standards on the reliability of the information presented, to auditing standard-setters and regulators who have a responsibility to clarify auditors' responsibility for misstatement in disaggregated numbers, and to audit firms that must provide guidance to ensure consensus in their auditors' judgments. Standard-setters should also consider the fact the FASB has also been considering issues related to balance sheet aggregation or netting of balances. As a consequence, the importance of the effects of aggregation on auditors' materiality judgments may be broader than the focus of the current study.

The implications of our results for audit firms, standard-setting, and regulation will also depend on the answers to a number of related questions. The first important question is what drives management's disaggregation decisions: attempts to inform or attempts to obscure important information. As we noted above, Heitzman et al. (2010) and Riedl and Srinivasan (2010) find that disaggregation is more likely when the information is more relevant to investors. However, other papers suggest substantial attempts to manage earnings or obscure important information. If auditors cannot determine the motive for disaggregation decisions, then it is quite possible that reliance on disaggregated numbers as materiality benchmarks will not result in the reliability gains our results suggest. This suggests that effective monitoring of disaggregation requires both

reporting and auditing standards that take into consideration this possibility. Recent standard-setting efforts related to the auditing of management disclosures are a first step in this direction. Although not the focus of this paper, future research could examine how managers might use disaggregation in a strategic fashion to conceal errors or misstatements, as well as how auditors monitor these possibilities.

Second, the disaggregated amounts provided in our experiment are voluntary disclosures. As a consequence, our findings suggest how auditors respond to the existing variation within U.S. GAAP or IFRS. For example, the recent SEC (2011a) study found that approximately half of IFRS companies report expenses by nature and half by function, and about one-third of those disaggregating based on function disclose no additional information based on nature. Our results should be directly relevant to these circumstances. But their relevance to a firmer disaggregation requirement is less clear. If IFRS or the Staff Draft recommendations are adopted in the U.S., then the new requirement could either increase or decrease the effects discovered in our experiment. On the one hand, those auditors who inferred relevance to investors based on the client's choice to provide the disaggregated numbers could judge the required numbers to be less relevant, and not consider them to be relevant materiality benchmarks. On the other hand, the IFRS standard requires companies to disaggregate only if the disaggregated amounts are relevant, and numbers provided as a result of a new accounting standard could be seen as a likely focus of SEC scrutiny. These factors could increase the perceived relevance of mandated disaggregated numbers as materiality benchmarks. While we speculate that these latter effects will dominate, the effects of a disaggregation requirement is an important topic for future research. We believe that a study of U.K. auditors' judgments would shed some light on the effect of a firmer disaggregation requirement, but any effect would be confounded with significant differences in the legal and regulatory regimes as well as historical and cultural differences.

REFERENCES

Acito, A., J. Burks, and W. B. Johnson. 2009. Materiality decisions and the correction of accounting errors. *The Accounting Review* 84 (3): 659–688.

American Institute of Certified Public Accountants (AICPA). 2010. *Audit Risk and Materiality in Conducting an Audit.* AU Section 312. Durham, NC: AICPA.

Berger, P., and R. Hann. 2003. The impact of SFAS No. 131 on information and monitoring. *Journal of Accounting Research* 41 (2): 163–223.

Berger, P., and R. Hann. 2007. Segment profitability and the proprietary and agency costs of disclosure. *The Accounting Review* 82 (4): 869–906.

Bhojraj, S., P. Hribar, M. Picconi, and J. McInnis. 2009. Making sense of cents: An examination of firms that marginally miss or beat analyst forecasts. *The Journal of Finance* 64 (5): 2361–2388.

Bloomfield, R., F. Hodge, P. Hopkins, and K. Rennekamp. 2011. *Does Enhanced Disaggregation and Cohesive Classification of Financial Information Help Credit Analysts Identify Firms' Operating Structure?* Working paper, Cornell University.

Botosan, C., and M. Stanford. 2005. Managers' motives to withhold segment disclosures and the effect of SFAS No. 131 on analysts' information environment. *The Accounting Review* 80 (3): 751–771.

Cho, Y. J. 2011. *Segment Disclosures, Internal Capital Markets, and Firm Value: Evidence from SFAS No. 131.* Working paper, Cornell University.

Conover, W. J. 1999. *Practical Nonparametric Statistics.* New York, NY: Wiley.

D'Souza, J., K. Ramesh, and M. Shen. 2010. Disclosure of GAAP line items in earnings announcements. *Review of Accounting Studies* 15 (1): 179–219.

Einhorn, H. J. 1970. The use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin* 73: 221–230.

Financial Accounting Standards Board (FASB). 2008. *Preliminary Views on Financial Statement Presentation*. Norwalk, CT: FASB.

Financial Accounting Standards Board (FASB). 2010. *Staff Draft of an Exposure Draft on Financial Statement Presentation*. Norwalk, CT: FASB.

Financial Accounting Standards Board and International Accounting Standards Board (FASB/IASB). 2009. *Analyst Field Test Results*. Norwalk, CT: FASB/IASB.

Frederickson, J., F. Hodge, and J. Pratt. 2006. The evolution of stock option accounting: disclosure, voluntary recognition, mandated recognition and management disavowals. *The Accounting Review* 81 (5): 1073–1093.

Heitzman, S., C. Wasley, and J. Zimmerman. 2010. The joint effects of materiality thresholds and voluntary disclosure incentives on firms' disclosure decisions. *Journal of Accounting and Economics* 49 (1): 109–132.

Hirst, E., L. Koonce, and S. Venkataraman. 2007. How disaggregation enhances the credibility of management earnings forecasts. *Journal of Accounting Research* 45 (4): 811–837.

International Accounting and Auditing Standards Board (IAASB). 2010a. *Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements*. New York, NY: International Federation of Accountants (IFAC).

International Accounting and Auditing Standards Board (IAASB). 2010b. *Agenda Item 9-A: Disclosures—Issues and Working Group Proposals*. (September 2010). New York, NY: International Federation of Accountants (IFAC).

International Accounting Standards Board (IASB). 2009. *Presentation of Financial Statements*. International Accounting Standard No. 1. London, U.K.: IASB.

Joe, J., A. Wright, and S. Wright. 2011. The impact of client and misstatement characteristics on the disposition of proposed audit adjustments. *Auditing: A Journal of Practice & Theory* 30 (2): 103–124.

Keune, M., and K. Johnstone. 2009. Staff Accounting Bulletin No. 108 disclosures: Descriptive evidence from the revelation of accounting misstatements. *Accounting Horizons* 19 (1): 19–54.

Keune, M., and K. Johnstone. 2012. Materiality judgments and the resolution of detected misstatements: The role of managers, auditors and audit committees. *The Accounting Review* 87 (5): 1641–1677.

Libby, R. 1981. *Accounting and Human Information Processing: Theory and Applications*. Englewood Cliffs, NJ: Prentice Hall.

Libby, R., and N. Seybert. 2009. Behavioral studies of the effects of regulation on earnings management and accounting choice. In *Accounting, Organizations and Institutions*, edited by C. Chapman, D. Cooper, and P. Miller. Oxford, U.K.: Oxford University Press.

Libby, R., M. Nelson, and J. Hunton. 2006. Recognition v. disclosure, auditor tolerance for misstatement and the reliability of stock-compensation and lease information. *Journal of Accounting Research* 44 (3): 533–560.

Libby, R., R. Bloomfield, and M. Nelson. 2002. Experimental research in financial accounting. *Accounting, Organizations and Society* 27:775–810.

Messier, W., N. Martinov-Bennie, and A. Eilifsen. 2005. A review and integration of empirical research on materiality: Two decades later. *Auditing: A Journal of Practice & Theory* 24 (2): 153–187.

Nelson, M. 2004. A review of experimental and archival conflicts-of-interest research in auditing. In *Conflicts of Interest: Problems and Solutions in Law, Medicine, and Organizational Settings*, edited by D. A. Moore, D. M. Cain, G. Loewenstein, and M. H. Bazerman. Cambridge, U.K.: Cambridge University Press.

Nelson, M., and H.-T. Tan. 2005. Judgment and decision making research in auditing: A task, person and interpersonal perspective. *Auditing: A Journal of Practice & Theory* 24 (1): 73–87.

Nelson, M., J. Elliott, and R. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings management decisions. *The Accounting Review* 77: 175–202.

Nelson, M., S. Smith, and Z.-V. Palmrose. 2005. The effect of quantitative materiality approach on auditors' adjustment decisions. *The Accounting Review* 80 (3): 897–920.

Ng, T., and H.-T. Tan. 2003. Effects of authoritative guidance availability and audit committee effectiveness on auditors' judgments in an auditor-client negotiation context. *The Accounting Review* 78 (3): 801–818.

Payne, J. W. 1976. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance* 16: 366–387.

Public Company Accounting Oversight Board (PCAOB). 2010a. *Consideration of Materiality in Planning and Performing an Audit*. Auditing Standard No. 11. Washington, DC: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2010b. *Evaluating Audit Results*. Auditing Standard No. 14. Washington, DC: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2011. *Standing Advisory Group Meeting: Auditing Financial Statement Disclosures*. (March 24, 2011). Washington, DC: PCAOB.

Riedl, E., and S. Srinivasan. 2010. Signaling firm performance through financial statement presentation: An analysis using special items. *Contemporary Accounting Research* 27 (1): 289–332.

Scheffé, H. 1959. *The Analysis of Variance*. New York, NY: John Wiley & Sons, Inc.

Schipper, K. 2007. Required disclosures in financial reports. *The Accounting Review* 82 (2): 301–326.

Securities and Exchange Commission (SEC). 1999. *Materiality*. Staff Accounting Bulletin No. 99. Washington, DC: Government Printing Office.

Securities and Exchange Commission (SEC). 2006. *Financial Statements*. Staff Accounting Bulletin No. 108. Washington, DC: Government Printing Office.

Securities and Exchange Commission (SEC). 2011a. *An Analysis of IFRS in Practice*. A Securities and Exchange Commission Staff Paper November 16, 2011. Washington, DC: Government Printing Office.

Securities and Exchange Commission (SEC). 2011b. Staff Accounting Bulletin No. 114. Washington, D.C.: Government Printing Office.

American Accounting Association

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 667–705

American Accounting Association
DOI: 10.2308/accr-50330

Is There Life after the Complete Loss of Analyst Coverage?

Simona Mola
U.S. Securities and Exchange Commission

P. Raghavendra Rau
University of Cambridge

Ajay Khorana
Independent Author

ABSTRACT: This paper examines the value of sell-side analysts to covered firms by documenting the effects on firm performance and investor interest after a complete loss of analyst coverage for periods of at least one year. We find that analyst coverage adds value to a firm both because it reduces information asymmetries about the firm's future performance and because it maintains investor recognition for that firm's stock. After the introduction of regulations that curtailed the informational advantage of analysts in the early 2000s, the investor recognition role of analysts remains important. Firms that lose all analyst coverage continue to suffer a significant deterioration in bid-ask spreads, trading volumes, and institutional presence but do not show a significant difference in subsequent performance relative to covered peers. In addition, controlling for other factors, we find that firms that lose all analyst coverage for one year are significantly more likely to delist than their covered peers. Our results provide insight into the reasons why firms place so much importance on analyst coverage.

Keywords: *analyst coverage*; *investor recognition*; *delisting prediction*.

We thank two anonymous referees, Katrien Bosquet, Naveen Daniel, Dave Denis, Barb Dwyer, John Harry Evans III (senior editor), Anzhela Knyazeva, S. P. Kothari, Chen Lin, Alexander Ljungqvist, Tim Loughran, Ronald Masulis, John McConnell, Lalitha Naveen, Jay Ritter, Henri Servaes, Ann Sherman, Lakshmanan Shivakumar, Anil Shivdasani, Issouf Soumare, Brett Trueman, Beverly Walther (editor), Kent Womack, and seminar participants at Arizona State University, Helsinki School of Economics, University of Western Ontario, the U.S. Securities and Exchange Commission, the 2007 FMA Meeting, the 2008 EFA Meeting, the 2009 EFMA Meeting, the 2009 International Conference in Corporate Finance and Financial Institutions at the City University of Hong Kong, the 2009 National Investor Relations Institute Meeting in Denver, the 2010 FMA European Meeting, the 2010 "People & Money" Conference at DePaul University, and the 2010 NFA Conference for helpful comments.

Editor's note: Accepted by Beverly Walther.

Submitted: March 2010
Accepted: October 2012
Published Online: October 2012

I. INTRODUCTION

Firms value analyst coverage. CEOs spend time and resources attempting to obtain and maintain coverage from sell-side analysts, who typically work for a brokerage house and provide investment research to their employer's clients. Rajan and Servaes (1997) and Cliff and Denis (2004) show that firms pay for the extent and quality of analyst coverage by underpricing their initial public offerings (IPOs). Krigman et al. (2001) find that the prospect of gaining access to better analyst coverage motivates firms to switch underwriters between their IPOs and their subsequent seasoned equity offerings (SEOs). Moreover, as Bushee and Miller (2010) document, firms that lose analyst coverage sometimes hire investor relations professionals to pitch their business to analysts. Kirk (2011) shows that firms without analyst coverage buy paid-for research. Why do firms value analyst coverage so much?

This question is especially relevant because firms appear to use the availability of analyst coverage as an important factor in their decision to list on public exchanges. Weild and Kim (2009) document that, over the past two decades, the number of firms listed on public exchanges has dropped by 22 percent, from 6,943 firms in 1991 to 5,401 firms in 2008. They argue that the "great depression in listings" is partly due to analysts ceasing to cover small- and mid-capitalization stocks. Similarly, Mehran and Peristiani (2010) argue that one reason behind the decision of recent IPO firms to abandon their public status and go private is that being listed on public exchanges no longer guarantees analyst coverage.

In this paper, we examine the value of analyst coverage to a firm by documenting the effects on investor interest, operating performance, and stock performance after a complete loss of analyst coverage. Specifically, we investigate two non-mutually exclusive hypotheses. The *private information* hypothesis argues that analysts have access to management's private information about a firm's prospects and convey it to investors through their coverage choices. McNichols and O'Brien (1997) argue that analysts drop coverage after receiving negative information, which is not reflected in their last outstanding forecast. According to the private information hypothesis, therefore, the loss of coverage will be followed by a decrease in firm performance. The *investor recognition* hypothesis argues that analysts direct investor attention to particular stocks. The loss of analyst coverage for a stock will reduce the investor interest for that stock, with adverse effects in liquidity, trading volume, and institutional presence (Merton 1987).

Our sample period ranges from 1984 to 2008. This is an appropriate period to analyze because it covers a series of significant analyst regulations and the enforcement of the Global Settlement (Bradshaw 2009). The analyst regulations, issued between 2000 and 2003, include Regulation Fair Disclosure (Reg FD), NASD Rule 2711, NYSE Rule 472, and Regulation Analyst Certification (Reg AC). It is likely that these rules have impacted analysts' access to private information and changed their economic incentives to cover firms. Consequently, the regulatory changes in the early 2000s represent a natural break in the analysis and enable us to draw inferences about why firms value analyst coverage.

We define a firm as having lost all analyst coverage when the firm, previously covered on I/B/E/S, receives no earnings forecasts during a calendar year.[1] Our sample consists of 16,662 observations of firms that receive no analyst coverage in a given year. These sample firm-year observations represent 19 percent of all U.S. firms covered by I/B/E/S from 1984 through 2008.

[1] This definition allows us to select a sample free from look-ahead bias. Such a bias arises if *ex post* information is used to establish termination of coverage. Instead, in our study, we infer the complete loss of coverage for a firm based only on the information available up to that date, given that analysts can choose to discontinue coverage in one year but then resume it in subsequent years.

Most sample firms are small, seasoned companies that have been previously covered for three years.

To address potential endogeneity issues in the data, we compare our sample firms to control firms that are similar to the sample firms across all dimensions, except for analysts' coverage decisions. Our research design thus consists of three steps. First, building on prior theoretical and empirical literature, we model an analyst's decision to not cover a firm in a given year t, as a function of the firm's performance and its potential for investor interest in year $t-1$. Investor interest is likely to be important to the analyst since it is related to the firm's ability to generate both trading and underwriting revenue for the analyst's employer. Second, we use a propensity score matching (PSM) technique to identify, for each sample firm that loses analyst coverage in a year, a covered firm that matches the sample firm on both factors. Finally, we measure the difference-in-differences effects after the loss of coverage over the next five years.

Before the introduction of the regulations that curtailed the informational advantages of analysts, we find that the effects of losing all analyst coverage are significantly related both to the information-gathering and processing abilities of analysts and to their ability to direct investor attention to stocks. The sample firms that lose analyst coverage before the regulatory changes experience a significant decline in both performance and investor interest indicators relative to their control firms. After the regulatory changes, the pattern shifts strikingly. The sample firms that lose analyst coverage after the regulatory changes continue to suffer a significant deterioration of bid-ask spreads, trading volumes, and institutional presence. Their performance indicators, however, are in line with covered peers, implying that analysts' coverage decisions no longer convey private information on firm performance.

After losing analyst coverage, our sample firms delist at a significantly higher rate than their control firms. A Cox proportional hazards model shows that, after controlling for a firm's performance and investor interest indicators, the loss of analyst coverage in year t significantly predicts the firm's delisting over the next ten years. This result is also economically significant: a firm that remains an orphan for one year is 11 percent more likely to delist in the next ten years than its control firm. The predictive power of the loss of analyst coverage for delisting holds both before and after the new regulations, although the economic magnitude shrinks in the period following the regulatory changes. Finally, a number of additional tests show that these results are robust and unlikely to be driven by endogeneity issues.

The paper makes several contributions to the literature on the value of analyst coverage. We show that analysts add value to firms both by gathering and processing information and by directing investor attention. Regulatory changes that reduce the information-gathering advantages of analysts change how analysts affect firm value. The investor recognition role of analysts, however, remains important. Lehavy and Sloan (2008, 359) argue that an unanswered puzzle in the literature is: "What factors cause investors to be cognizant of some securities, but not others?" Our results suggest that analysts are an important factor in bringing covered stocks to the attention of investors. Consistent with the predictions in Merton (1987), we find that the loss of analyst coverage reduces the number of investors who recognize and trade the firm's stock, with attendant effects of widening the bid-ask spread and decreasing trading volume and institutional presence. We also contribute to prior literature by analyzing the relation between analyst coverage and delisting. After controlling for firm performance and investor interest characteristics, the complete loss of analyst coverage significantly predicts a firm's delisting. Our findings thus provide empirical evidence for the arguments made by Weild and Kim (2009) and Mehran and Peristiani (2010) about the scarcity of analyst coverage as a reason for the decrease (increase) in U.S. listings (delistings).

Overall, our results suggest that managers are correct to worry about a complete loss in analyst coverage. Given that the demand curve for a stock with low investor recognition is downward-sloping, a demand shock can significantly depress its stock price, making the firm

vulnerable to possible acquisitions or other delisting events. Moreover, firms with low investor recognition face difficulty raising capital and investing in positive-NPV projects (Lehavy and Sloan 2008). When firms remain without analyst coverage for several years, finance and investment projects are likely to become exceedingly expensive, forcing firms to forgo value-enhancing opportunities and ultimately jeopardizing their growth and survival prospects. Without coverage, firms are likely to disappear from investors' radar and their odds of survival in the financial markets drop significantly. To reverse the philosopher Bishop Berkeley's dictum, for these firms, *percipi est esse*, to be perceived is to be.

Section II next develops the hypotheses and reviews the related literature, while Section III describes the regulatory changes over our sample period. In Section IV, we discuss data and sample descriptive statistics, and model an analyst's decision to provide coverage on a firm. Section V reports the effects on firm performance and investor interest characteristics after losing coverage, while Section VI reports the effects after regaining coverage. We discuss robustness checks in Section VII and conclude in Section VIII.

II. HYPOTHESES DEVELOPMENT AND RELATED LITERATURE

We investigate two non-mutually exclusive hypotheses as to how analysts impact firm value. The *private information* hypothesis argues that analysts add value by conveying private information from a firm's management to investors. Brennan and Subrahmanyam (1995) and Ellul and Panayides (2009), among others, argue that analyst research is valuable because it reduces informational asymmetries between investors and a firm's insiders. Lehavy et al. (2011) find that the informativeness of analyst reports is greater for firms with less readable 10-K filings. Frankel et al. (2006) show that analyst reports are more informative when the potential brokerage profits are higher (e.g., high trading volume, high volatility, and high institutional ownership) but less informative when information-processing costs are higher (e.g., in firms with more business segments). They also argue that the informativeness of analyst research and the informativeness of financial statements are complementary.

However, analysts do not necessarily add value only by serving as a conduit for information privately obtained from management. The *investor recognition* hypothesis argues that analysts make investors aware of stocks. Analysts release common industry-level information (Pietroski and Roulstone 2004), piggyback on corporate news (Altınkılıç and Hansen 2009), or boost the outlook of clients' stock (James and Karceski 2006). Easley et al. (1998) argue that when analysts release a report on a stock, they add no private information but generally repackage publicly available information. In doing so, they attract investors' limited attention by showcasing one stock among many. This argument is consistent with Merton (1987) in that analyst coverage broadens the investor recognition of firms and improves the liquidity of the firms' stock (Baker et al. 2002; Irvine 2003).

What distinguishes the two hypotheses is the existence of material private information about a firm. When analysts have private information about a deterioration of the firm's growth prospects, they can decide to discontinue coverage on that firm, rather than issue a negative recommendation. McNichols and O'Brien (1997) argue that analysts are reluctant to issue negative information because they either fear jeopardizing potential client relationships or fear losing access to management. This implies that the negative information is not reflected in their last forecast. The private information hypothesis predicts that a drop in coverage driven by private information about a firm's outlook is likely to be followed by a decline in that firm's performance. In contrast, a drop in coverage driven by reasons unrelated to private information about a firm's performance (e.g., the firm is considered unable to generate sufficient underwriting and/or trading revenue) or exogenous factors (e.g., analyst's resignation) is unlikely to be followed by a decline in firm performance. The

investor recognition hypothesis predicts that the adverse effects for that firm will result in lower stock liquidity, lower contemporaneous returns, and higher expected returns to compensate shareholders for being imperfectly diversified (Amihud and Mendelson 1986; Merton 1987).[2]

The binary distinction between the two hypotheses is somewhat artificial because the extent of "private information" and "investor recognition" can vary significantly from one stock to another. For example, analysts can convey their own private information, as opposed to management's private information, by producing forecasts or recommendations from public data, highlighting public nonfinancial data in reports to clients, or making coverage decisions. These activities have the potential to play an information role beyond making firms visible to investors.

Two studies examine the price effects of partial decreases in analyst following and find conflicting results. Kelly and Ljungqvist (2011) investigate coverage reductions that are the result of brokerage firms downsizing their research operations in the early 2000s. They argue that reductions in the supply of information have a direct adverse effect on asset prices. Cumulative excess returns average –0.47 percent on the day of a coverage termination. The price impact is reduced when other analysts continue to cover the stock but the observed price reaction is not mean-reverting over the first month of trading or over the next six to 24 months. In contrast, Kecskés and Womack (2007) document that investors initially overreact to a decrease in the number of analysts covering a stock. In their sample, a drop in analyst following in one year produces negative returns that are followed by higher positive returns the next year.

Our paper differs from these studies across several dimensions. First, we analyze the effects on firm performance, investor interest, and delisting rates after the loss of analyst coverage, and relate them to analyst information-gathering and processing abilities. Second, our analysis of complete losses of analyst coverage allows for a clearer test than partial reductions in coverage because it is likely that a firm continues to be visible to investors when it experiences a partial loss in analyst coverage. Finally, the extent of our sample period from 1984 up to the end of 2008 offers the opportunity to examine the effects of losing analyst coverage after the regulatory changes in the analyst information environment.

III. REGULATORY CHANGES

Our 1984–2008 sample period covers a series of important regulatory changes that affected analyst activities. Adopted by the Securities and Exchange Commission (SEC) in August 2000, Reg FD states that managers cannot privately disclose material information to financial analysts (Release No. 33-7881). Gintschel and Markov (2004) show that, consistent with Reg FD curtailing the flow of information from managers to analysts, the average price impact associated with the dissemination of analysts' information drops by 28 percent from the pre-Reg FD level.

In May 2002, the SEC approved similar changes to both National Association of Securities Dealers (NASD) Rule 2711 and NYSE Rule 472 (Release No. 34-45908). NASD Rule 2711 now requires that every brokerage firm discloses in its research reports the distribution of stock ratings across its coverage universe. Barber et al. (2006) document that, as a result of this rule, the percentage of buys has decreased steadily since mid-2000. Rule 2711 also prohibits NASD members from tying analysts' compensation to their employers' investment banking transactions and from offering favorable research, rating, or price target to a firm as inducement for future business. Chen and Chen (2009) find that the implementation of this rule has enhanced analysts'

[2] Our analysis depends on identifying publicly available variables that play a role in an analyst's decision to cover a firm. To the extent that analysts use private information in their coverage choices, our tests may suffer from the joint hypothesis problem that is common to all event studies, and we cannot conclusively detect causality. However, matching on observable characteristics greatly reduces this concern, and our additional tests on exogenous losses of coverage further alleviate this issue.

independence. Another relevant provision of NASD Rule 2711, stated in paragraph (f)(5), requires analysts to issue a final research report prior to terminating coverage of a stock so that investors receive the benefit of the guidance from the analysts who recommended the stock in the first place.[3] Finally, the SEC approved Reg AC in February 2003 (Release No. 33-8193), a few months before the Global Settlement was reached between the SEC, NASD, NYSE, and ten large investment banks. Among other features, Reg AC requires a statement from analysts about their compensation.[4]

The enactment of all these regulations over a short period makes it virtually impossible to analyze the incremental impact of the loss of coverage on firms from any one regulation. However, the relatively short enactment period of these rules forms a natural break in which to examine the change in the effects of losing analyst coverage prior to 2000 and after 2003. These regulations likely changed the information environment for analysts and the incentives to cover firms. Reg FD is likely to have reduced the informational advantage of analysts, while NASD Rule 2711, NYSE Rule 472, and Reg AC are likely to have curbed analysts' incentives to gain underwriting business potentially through biased coverage. In the wake of these regulatory changes, analysts may be less likely to drop coverage on firms based on private information about the firms' future performance, and they may be less likely to drop coverage on firms that generate little or no underwriting business to their employers.

IV. DATA AND SAMPLE DESCRIPTIVE STATISTICS

We define a firm as having lost all analyst coverage when analysts have issued no earnings forecasts on the firm during a calendar year.[5] From all firms covered on the I/B/E/S detail files, we isolate firms that receive no analyst coverage during at least one calendar year, from 1984 through 2008. To construct our sample, we apply three screening criteria to this population of firms. First, we remove American Depository Receipts (ADRs) and retain only firms incorporated in the U.S. because foreign firms may be covered by other analysts who we cannot track. Second, sample firms must have ordinary common shares publicly traded on the NYSE, AMEX, or NASDAQ. We remove certificates, shares of beneficial interest, depository units, units of beneficial interest, units of limited partnership interests, depository receipts, real estate investment trusts (REITs), and closed-end funds. Third, using CRSP delisting dates, we retain those firms that are still publicly traded at the end of the year when analyst coverage is absent. This eliminates firms that lose coverage in a year only because they have been liquidated, acquired, or delisted for other reasons in that year. Our final sample consists of 16,662 firm-year observations of U.S. firms that are still publicly traded at the end of the year when they receive no analyst coverage. Most firms in our

---

[3] In Section VII, our analysis of a large sample of the final reports issued in compliance with NASD Rule 2711 (f)(5) suggests that these reports offer little guidance to investors.

[4] According to the text of Reg AC (part I, section A, point (B)), analysts need to include in their research reports "a statement by the research analyst (or analysts) certifying either: (1) that no part of his or her compensation was, is, or will be directly or indirectly related to the specific recommendations or views contained in the research report; or (2) that part or all of his or her compensation was, is, or will be directly or indirectly related to the specific recommendations or views contained in the research report. If the analyst's compensation was, is, or will be directly or indirectly related to the specific recommendations or views contained in the research report, the statement must include the source, amount, and purpose of such compensation, and further disclose that it may influence the recommendation in the research report."

[5] In our research design, we use calendar year in preference to fiscal year-end and make no assumptions on the length of time a firm can have outstanding coverage before it is considered stale. This simplifies the propensity score matching but may affect the firms (37 percent of our sample) whose fiscal year does not end in December. In robustness checks, we find that our results are qualitatively unaffected by the use of calendar year.

sample are listed on NASDAQ (71 percent) and are not recent IPOs. Fewer than 11 percent of these firms have been listed for less than three years. The median trading age is about eight years.

The data used in this paper come from multiple sources. Stock returns, bid-ask spreads, and trading volume data are obtained from CRSP, accounting data from Compustat, institutional holdings from CDA/Spectrum Institutional 13f Holdings, and analyst recommendations and earnings forecasts from I/B/E/S. To classify analysts' employers, we use firm websites, Factiva, Businessweek, and Nelson's Research Directory.

Figure 1 shows the evolution of sample firms by year from 1984 to 2008, and compares the number of previously covered firms that receive no analyst coverage in a given year to the number of firms that analysts actually cover in that year. We apply the same criteria used to screen our sample firms to the universe of covered firms on I/B/E/S. The covered firms are thus U.S. firms whose ordinary shares are publicly traded on the main domestic exchanges at the end of the year when they receive analyst coverage.

Over the 25-year period, the number of firms that lose coverage appears to rise and fall with the business cycle, with more firms being dropped by analysts during economic contractions and fewer firms being dropped during economic booms. Since our sample is restricted to firms that continue to be traded publicly at the end of the year when coverage is absent, the loss of coverage in this period is not simply due to firms failing or being delisted. It may seem plausible that more analysts leave the business during contractions than during booms, resulting in a mechanical relation between business cycle and coverage losses. However, Figure 1 shows that the number of analysts on I/B/E/S has stayed relatively stable in the 2000s. An alternative hypothesis is that, during contractions, analysts are more likely to work shorter hours, selectively paying more attention to those firms that generate significant fees, while dropping other firms from their coverage universe without replacement.[6]

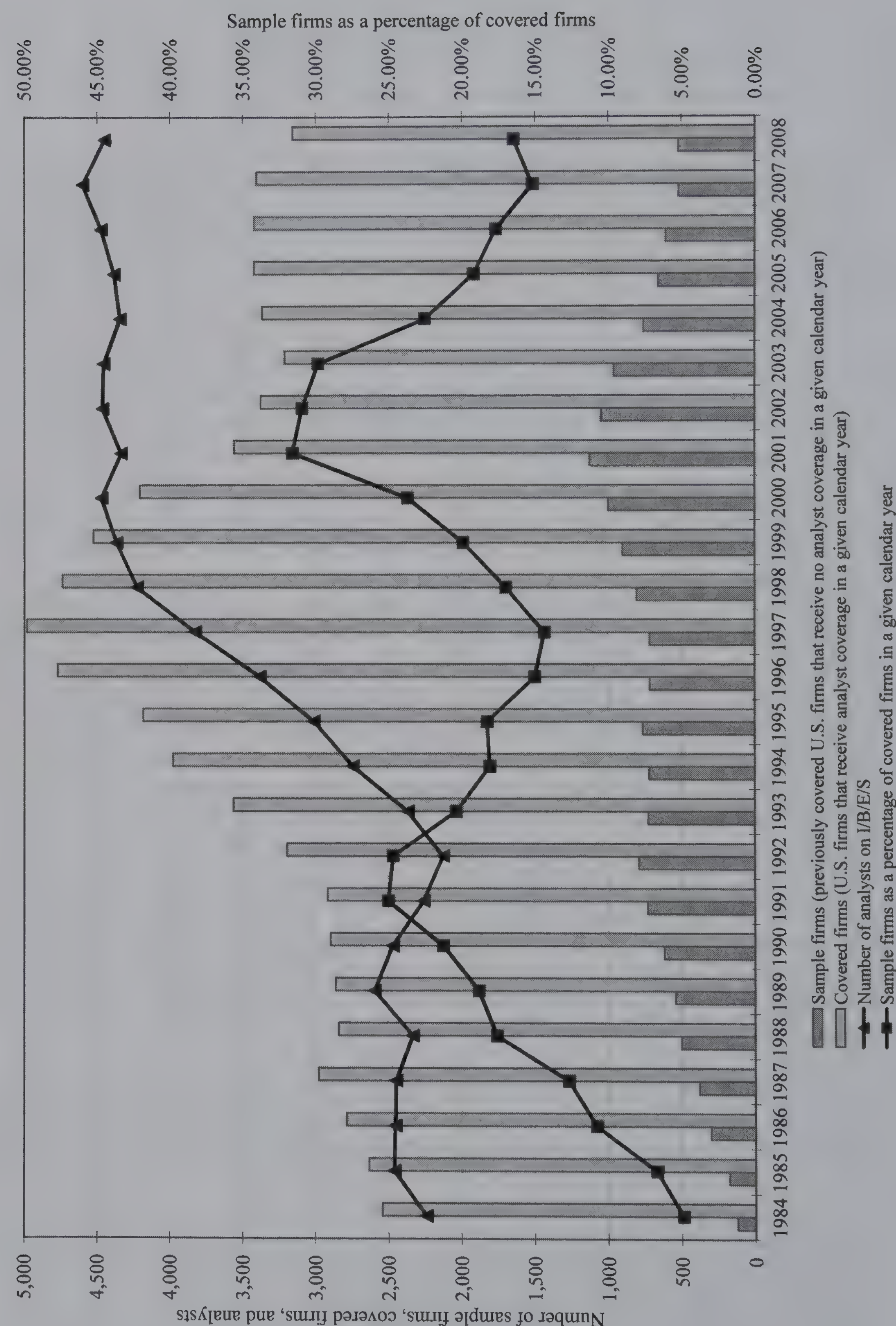
Panel A of Table 1 categorizes our sample of 16,662 firm-year observations into three subperiods that coincide with different regulatory environments in which analysts operated from 1984 to 2008: (1) the pre-Reg FD subperiod (1984–2000); (2) the post-Reg FD subperiod (2001–2002) when most changes in analyst regulation occurred; and (3) the post-Reg AC subperiod (2003–2008). Before the regulatory change subperiod, an average of 617 firms lost coverage every year with each analyst covering an average of 13 firms. In contrast, after the regulatory change subperiod, 669 firms, on average, lose coverage annually while analysts cover an average of 10 firms each. Starting from 2001 therefore, more analysts have concentrated their attention on fewer firms.

In Panel B of Table 1, we sort the 16,662 sample observations by the number of consecutive years with and without analyst coverage. First, we determine the length of time that a sample firm has been covered before becoming an orphan as a fraction of years from the day of the first earnings estimate to the day of the last earnings estimate by any analyst on I/B/E/S. About 21 percent of sample firms fell out of analyst favor after only one year of coverage, while about 21 percent received continuing coverage for more than seven years. The median duration of analyst coverage for our treatment sample firms is three years. Second, at the end of a given calendar year, we categorize each firm-year observation by the number of consecutive years spent without analyst coverage. The duration of the loss of analyst coverage is negatively related to the number of years of coverage prior to the loss ($\rho = -0.06$). *Ex post*, the median sample firm remains an orphan for

---

[6] Using the first two digits of the North American Industry Classification System (NAICS) to classify firms by industry, we test the hypothesis that sample firms and covered firms are from populations with the same distribution by year and industry. We find no significant difference in the distributions at the conventional levels, according to two-sample two-tailed Wilcoxon rank-sum tests. A large proportion of our sample firms are manufacturing (41 percent) or financial firms (17 percent), and they do not cluster in time. We find significantly different distributions only for the firms operating in the high-tech and Internet industry (p-value = 0.07). As defined in Loughran and Ritter (2004), these firms cluster around the end of the dotcom bubble but account for only 10 percent of the sample.

FIGURE 1
Sample by Year



(continued on next page)

FIGURE 1 (continued)

The figure plots the evolution in the number of sample firms, from 1984 to 2008. "Sample firms" are previously covered U.S. firms that receive no earnings estimates during a calendar year. "Covered firms" are U.S. firms that receive at least one analyst estimate during a calendar year. To be included in the analysis, the common stock of sample firms and covered firms must still be publicly traded on the main domestic exchanges (i.e., the New York Stock Exchange, American Exchange, and NASDAQ) at the end of a calendar year. The figure also displays the number of analysts on I/B/E/S and the number of sample firms as a percentage of covered firms in a given calendar year.
Data come from I/B/E/S Detail and CRSP databases.

three years before regaining coverage or being delisted. *Ex ante*, the distribution of our sample is monotonically decreasing with 4,836 firms receiving no coverage for the first year, 3,040 of these firms remaining uncovered for the second year in a row, while others regain analyst coverage or delist, and so on. This procedure is free from any look-ahead bias. At the end of each calendar year, market participants learn about the lack of analyst coverage on a firm and the duration of the coverage loss for the firm, up to that date.

Do Firms Lose Analyst Coverage Abruptly?

In Table 2, we report analyst earnings per share (EPS) estimates and recommendations on sample firms in the five years prior to year t, which is the first year when firms lose all analyst coverage. Since the level of the EPS is a firm-specific arbitrary number unrelated to the firm's economic performance, we focus on changes in EPS. The mean EPS estimate declines steadily from \$0.58 in year $t-5$ to \$0.13 in year $t-1$. The annual changes are significant at the 1 percent level. Similarly, the mean industry-adjusted EPS estimate is $-$\$0.59 in year $t-5$ and drops significantly to $-$\$1.05 below the industry estimate in year $t-1$. The fact that the mean industry-adjusted estimate becomes increasingly negative over time indicates that analysts believe that sample firms are performing, on average, worse than the industry, not that the industry itself is out of favor. Recommendations (with strong buy = 1 and sell = 5) slowly worsen for the average sample firm from 2.05 in year $t-5$ to 2.42 in year $t-1$. Industry-adjusted recommendations show the same pattern.

The number of annual estimates for the average sample firm gradually decreases from 10 to 3 in the five years examined with a commensurate decrease in the number of analysts covering the firm. The quality of analyst coverage declines over time with mean analyst forecasting ability, which is proxied by the proportional mean absolute forecast error (PMAFE) for an analyst in a year (Clement 1999). It worsens sharply in year $t-1$, suggesting that superior analysts decide to stop coverage earlier. Both the proportion of Institutional Investor star analysts and the proportion of analysts affiliated with investment banks decrease over the five-year period.

In contrast, the proportions of analysts employed by independent brokers or paid-for research firms gradually increase. In our data, 352 distinct firms purchase coverage from 15 paid-for research firms, such as Taglich Brothers Inc., J. M. Dutton & Associates, and Spelman Financial. Overall, Table 2 indicates that losing coverage is not a precipitous process that is triggered by some external shock; instead, sample firms experience a gradual loss of analyst interest.

Which Firms Lose Coverage?

We model the analyst's decision to drop coverage on a firm as a function of the firm's operating and stock performance, and its potential for investor interest. McNichols and O'Brien (1997) argue that analysts prefer to report good news than bad news because good news is easier to sell to a broader audience. This implies that, instead of downgrading, analysts are likely to drop

TABLE 1

Sample by Subperiods and Number of Years with (without) Analyst Coverage

Panel A: Subperiods

Subperiod	Sample Firm-Year Observations	Average Number of Sample Firms Per Year	Average Number of Covered Firms Per Year	Average Number of Analysts Per Year	Average Number of Covered Firms Per Analyst
Pre-Reg FD, 1984–2000	10,482	617	3,565	2,928	13.23
Post-Reg FD, 2001–2002	2,168	1,084	3,468	4,400	8.95
Post-Reg AC, 2003–2008	4,012	669	3,331	4,451	9.63
All sample period	16,662	666	3,501	3,411	12.02

Panel B: Number of Years with (without) Analyst Coverage

Number of Years with Analyst Coverage	Sample Firm-Year Observations	As % of All Sample Firm-Year Observations	Number of Years without Analyst Coverage	Sample Firm-Year Observations	As % of All Sample Firm-Year Observations
$\leq$ 1 year	3,571	21.43%	1st year	4,836	29.02%
1 to 3 years	4,737	28.43%	2nd year	3,040	18.25%
3 to 7 years	4,851	29.11%	3rd year	2,168	13.01%
> 7 years	3,503	21.03%	4th year and beyond	6,618	39.72%
Total	16,662	100.00%	Total	16,662	100.00%

Panel A categorizes the 16,662 sample firm-year observations by subperiods that coincide with different regulatory environments in which analysts operated from 1984 through 2008. Also, Panel A reports the average number of sample firms, covered firms, and analysts per year by subperiods. In Panel B, sample observations are categorized by the number of subsequent years during which sample firms consistently received analyst coverage. The duration of analyst coverage for a firm is defined as the fraction of years between the day of the first report and the day of the last report on that firm by any analyst on I/B/E/S. Sample observations are also categorized by the number of subsequent years during which sample firms receive no analyst coverage. To illustrate the methodology, consider a sample firm *i* that loses coverage in 1994 for the first year and regains it in 1997. *Ex ante*, market participants can only observe the number of years spent without coverage at the end of every year *t*, sequentially, year by year. Our sample contains three firm-year observations of firm *i*: *i*-1994, *i*-1995, and *i*-1996. As we sort these observations by the number of years without analyst coverage, observation *i*-1994 is categorized as "1st year without analyst coverage"; observation *i*-1995 as "2nd year without analyst coverage"; and observation *i*-1996 as "3rd year without analyst coverage."
Data come from I/B/E/S Detail and CRSP databases.

coverage of firms that are performing poorly or that they believe will perform poorly. Similarly, factors that promote investor interest are also likely to generate revenue for the analyst's employer, either trading (Irvine 2001, 2004) or underwriting revenue (Clarke et al. 2007). Analysts may potentially consider dropping coverage of firms that are unable to generate these types of revenue.

The performance indicators include sales growth rate, operating cash flow on assets, and return on assets (ROA). Following Zmijewski (1984), we compute total liabilities/total assets as a predictor of financial distress. We also calculate Ohlson's O-score (Ohlson 1980) as a predictor of bankruptcy with higher values indicating greater distress risk. Shumway (2001) finds that market-driven indicators better predict bankruptcy events than accounting ratios, and we therefore include excess

TABLE 2

Analyst Earnings Estimates and Recommendations prior to the First Year without Analyst Coverage

	Levels			Changes	
Year *t*–*n*	***t*–5**	***t*–3**	***t*–1**	***t*–5 to *t*–1**	***t*–3 to *t*–1**
EPS estimate	0.58	0.37	0.13	−0.47	−0.33
Industry-adjusted EPS estimate	−0.59	−0.82	−1.05	−0.43	−0.29
Recommendation	2.05	2.14	2.42	0.72	0.65
Industry-adjusted recommendation	−0.10	−0.01†	0.23	−0.10	−0.01
Number of estimates	10	8	3	−6	−4
Number of analysts	4	3	2	−2	−1
Analyst forecasting ability, PMAFE (%)	−1.38	−1.37	2.13	4.17	3.65
Proportion of star analysts (%)	9.31	8.37	6.36	−0.44	−0.20†
Proportion of analysts affiliated with					
investment banks (%)	90.89	90.11	86.15	−3.13	−2.25
independent brokers (%)	8.40	8.94	11.70	3.07	1.77
paid-for research firms (%)	0.71	0.95	2.15	0.32†	0.28†

All mean values are different from zero at the 1 percent level according to the standard one-sample, two-tailed t-test, except for the † values, which are not significantly different from zero at the conventional levels (p-value $>$ 0.10).

Mean analyst earnings-per-share (EPS) estimates and mean recommendations for sample firms are determined in the five-year period prior to the first year without analyst coverage. Industry-adjusted EPS estimate is the mean difference between sample EPS estimate and the mean industry EPS estimate. Industry is defined using the two-digit NAICS code. Recommendations range from 1 = strong buy to 5 = sell. Industry-adjusted recommendation is the mean difference between the sample recommendation and the mean industry recommendation. The EPS estimates come from the I/B/E/S Detail tapes, while recommendations come from the I/B/E/S Recommendation tapes. Earnings estimates are available on the I/B/E/S tapes starting 1981, while analyst recommendations are available starting 1994. Number of estimates is the mean number of EPS estimates issued on sample firms in year *t*–*n*. Number of analysts is the mean number of analysts issuing at least one report on the sample firms in year *t*–*n*. Analyst forecasting ability is determined as the average proportional mean absolute forecast error (PMAFE) across all the stocks covered by an analyst *i* in a given year. Following Clement (1999), the PMAFE is defined as the percent deviation of the absolute forecast error for analyst *i* on stock *j* in a year from the mean absolute forecast error made by all analysts covering stock *j* in that year. Negative (positive) values of the average PMAFE represent better (worse) than average performance. Star analysts are the analysts identified once a year by Institutional Investor as members of the All-American Research Team. Firms that employ analysts are assigned to one of the following categories: (1) investment bank, if analysts' employer is affiliated with an investment bank; (2) independent broker, if analysts' employer has no investment banking affiliation and provides research that is tied to brokerage services and/or institutional trading; or (3) paid-for research firm, if analysts' employer provides research that is directly or indirectly paid by covered firms.

return and idiosyncratic volatility. Finally, since the performance indicators are correlated, we report a composite performance index as the first principal component of the performance variables.

Our measures of potential investor interest and consequent revenue generation for the analyst's employer include market capitalization, book-to-market (B/M) ratio, trading volume, bid-ask spread, total institutional holdings, number of institutions, and the number of M&A deals and equity issues in which the firm was involved in the last three years. Collins et al. (1987) note that firm size can serve as a proxy for the potential to generate brokerage revenue, since larger firms are more visible, raise larger amounts of capital, have more traded shares, and engage in larger acquisitions. Jegadeesh et al. (2004) argue that analysts have significant economic incentives to endorse glamour firms with low book-to-market ratios, because these firms tend to be widely held by institutional clients. As in Bhushan (1989), we compute institutional holdings and the number of

institutions holding a stock as a measure of buy-side interest in that stock. We also determine a composite index for investor interest. Appendix A presents more information about how we compute each proxy.

Table 3 reports median and mean performance indicators and investor interest characteristics for sample and covered firms in year $t-1$. All variables are winsorized at the 1 percent and 99 percent levels. Given that the loss of analyst coverage is defined on an annual basis, year t marks the year when the firm is not covered by any analyst. More specifically, year t can denote the first year the firm spends without analyst coverage, the second year, and so on. We focus on median values, although our inferences on means are qualitatively similar. All median (mean) differences between sample and covered firms in the table are significantly different from zero at the two-tailed 1 percent level, using a two-sample Wilcoxon rank-sum test (two-sample t-test).

In the year prior to losing analyst coverage, the median sample firm is significantly less profitable and less financially stable than the median firm covered on I/B/E/S. The median composite performance index is –20.11, compared to 19.41 for covered firms. The individual factors that compose the performance index are also significantly worse than covered firms. Sample stocks are "losers" in terms of stock performance, earning significantly lower excess returns with greater

TABLE 3

Performance Indicators and Investor Interest Characteristics for Sample Firms and Covered Firms

	Sample Firm-Year Observations			Covered Firm-Year Observations		
	Median	**Mean**	**n**	**Median**	**Mean**	**n**
Performance Indicators						
Sales growth (%)	0.84	10.38	14,292	5.96	11.56	76,347
Cash flow/total assets (%)	2.53	−5.18	14,482	7.45	4.14	75,625
ROA (%)	0.51	−10.61	14,647	3.35	−0.28	78,779
Total liabilities/total assets	0.55	0.56	14,652	0.53	0.53	78,615
Ohlson's O-score	0.33	0.42	14,478	−1.04	−1.03	77,606
Excess return (%)	−6.39	−3.38	14,732	2.08	4.90	80,237
Stock volatility (%)	12.69	15.20	14,689	9.35	11.20	79,278
Performance index (%)	−20.11	−40.57	13,599	19.41	11.29	72,131
Investor Interest Characteristics						
Market capitalization ($ millions)	27.91	104.40	14,782	243.97	1,566.65	79,786
B/M ratio	0.78	0.86	14,569	0.51	0.59	78,769
Trading volume (millions of shares)	1.75	8.85	14,744	11.19	57.25	80,251
Bid-ask spread (%)	4.55	6.11	12,983	1.69	2.45	67,574
Total institutional holdings (%)	11.41	17.24	15,115	37.54	40.61	80,401
Number of institutions	9	13.41	15,052	40	82.65	80,171
Number of M&A deals	0	0.80	16,662	1	1.83	87,516
Number of issues	0	0.05	16,662	0	0.25	87,516
Investor interest index	−0.72	−0.74	12,387	0.13	0.13	60,909

Year t marks the calendar year during which sample firms receive no analyst coverage. All performance indicators and investor interest characteristics are determined at the end of year $t-1$. Accounting data come from the Compustat Fundamentals Annual database. Variables are winsorized at the 1 percent and 99 percent levels. All median (mean) differences between sample and covered firms in the table are significantly different from zero at the two-tailed 1 percent level, using a two-sample Wilcoxon rank-sum test (two-sample t-test).

volatility than covered stocks. A comparison between the distributions of performance indicators for sample firms and the quintile breakpoints of performance indicators for covered firms shows that the sample firms are disproportionately represented in the worst quintiles of performance.

The sample firms are typically small-cap firms with higher median B/M ratios and bid-ask spreads, and lower trading volume than covered firms. Sample stocks are not heavily traded, with annual trading volume of 1.75 million shares for the median sample stock. Both institutional holdings and the number of institutions holding sample firms' stock are significantly lower. The median firm in our sample engages in no M&A activity. This is also true for the typical firm covered on I/B/E/S. The mean M&A and issue activity—both number of deals and dollar amounts—of sample firms is considerably lower than the mean activity of covered firms. The overall investor interest index is also significantly lower than for the median covered firm. Again, most investor interest characteristics of sample firms fall into the worst quintiles relative to covered firms.

We next use a logistic regression to model the probability that analysts decide not to cover a firm. In models 1 to 4 of Table 4, the dependent variable is a binary variable that is equal to 1 in year t when a firm loses analyst coverage, or 0 otherwise. The independent variables are the performance indicators and investor interest characteristics from Table 3. Year, industry, and exchange fixed effects are included in each regression. Model 1 suggests that, in the 1984–2008 sample period, the likelihood of losing analyst coverage is inversely related to sales growth, cash flow/total assets, ROA, and excess return. It is positively related to financial distress proxies, such as debt ratio, Ohlson's O-score, and idiosyncratic volatility. The investor interest characteristics also significantly explain the probability of losing analyst coverage: market capitalization, trading volume, institutional holdings, number of institutions, the number of M&A deals, and the number of issues are all negatively related, while B/M ratio and bid-ask spread are positively related to the probability of losing coverage. In untabulated regressions, both the composite performance index and the composite investor interest index are significantly negatively related to the probability of losing coverage.

In models 2 to 4, we split the 25-year sample period into the three regulatory subperiods. In the pre-Reg FD subperiod, all independent variables are significant predictors. In the post-Reg FD subperiod however, the probability of losing coverage is mostly driven by a firm's potential for investor interest. While the coefficients of sales growth and the Ohlson's O-score are significant in the post-Reg AC subperiod, it is mainly a firm's potential to generate brokerage business, rather than its performance, that may induce analysts to drop coverage.[7]

To identify the control firms for model 5, we use the PSM methodology (Villalonga 2004; Cooper et al. 2005). This matching method has the advantage of identifying a control group of firms screened along multiple dimensions. Propensity scores are computed by running the logistic regression model 1 of Table 4 over 15 independent variables (sales growth, cash flow/total assets, ROA, total liabilities/total assets, Ohlson's O-score, excess return, stock volatility, market capitalization, B/M ratio, trading volume, bid-ask spread, total institutional holdings, number of institutions, number of M&A deals, and number of issues) computed in year $t-1$, along with year, industry, and exchange fixed effects. Using a one-to-one nearest available neighbor matching within calipers defined by propensity scores (Rosenbaum and Rubin 1985), we find matches for 8,732 of the 16,662 sample firm-year observations. The inability to find a control firm is due to missing variable values or the lack of overlap between the propensity score ranges for sample and control firms.

To verify that our matching procedure results in a number of good matches, we run two-tailed Wilcoxon rank-sum tests on the equality between the median values of each explanatory variable

[7] When we run logit regressions with only the set of performance indicators across the three subperiods, Wald Chi-square tests significantly reject the hypothesis that all predicted coefficients are simultaneously equal to 0.

TABLE 4

Logistic Regression for the Probability that a Firm Will Lose Analyst Coverage

	Before Matching				After Matching
	All Sample Period (1)	Pre-Reg. FD, 1984–2000 (2)	Post-Reg. FD, 2001–2002 (3)	Post-Reg. AC, 2003–2008 (4)	All Sample Period (5)
Intercept	−0.44	−0.19	2.14**	1.63**	−0.19
	(−0.89)	(−0.38)	(2.09)	(2.22)	(−0.32)
Performance Indicators					
Sales growth$_{t-1}$	−0.12***	−0.12***	−0.07	−0.10***	0.02
	(−6.03)	(−4.82)	(−1.41)	(−2.82)	(0.94)
Cash flow/Total assets$_{t-1}$	−0.43**	−0.56*	−0.24	−0.41	0.04
	(−1.93)	(−1.90)	(−0.45)	(−0.95)	(0.16)
ROA$_{t-1}$	−0.72***	−1.09***	−0.81*	−0.30	0.10
	(−3.63)	(−3.92)	(−1.83)	(−0.79)	(0.48)
Total liabilities$_{t-1}$/Total assets$_{t-1}$	0.74***	1.07***	0.30	0.30*	0.05
	(8.90)	(9.38)	(1.42)	(1.79)	(0.49)
Ohlson's O-score$_{t-1}$	0.08***	0.11***	0.05*	0.06***	−0.01
	(7.02)	(6.96)	(1.67)	(2.59)	(−0.74)
Excess returns$_{t-1}$	−0.03	−0.10***	−0.03	−0.06	0.05
	(−1.17)	(−2.66)	(−0.40)	(−0.96)	(1.40)
Stock volatility$_{t-1}$	1.75***	2.76***	−0.20	0.55	−0.37
	(7.61)	(8.78)	(−0.36)	(1.12)	(−1.36)
Investor Interest Characteristics					
Ln(Market capitalization$_{t-1}$)	−0.64***	−0.55***	−0.81***	−0.80***	−0.04*
	(−29.04)	(−20.07)	(−12.66)	(−15.68)	(−1.80)
B/M ratio$_{t-1}$	0.17***	0.26***	0.01	0.11**	−0.01
	(7.99)	(8.88)	(0.15)	(2.23)	(−0.51)
Ln(1 + Trading volume$_{t-1}$)	−0.38***	−0.38***	−0.32***	−0.38***	0.01
	(−22.70)	(−16.53)	(−6.90)	(−11.84)	(0.69)
Bid-ask spread$_{t-1}$	3.96***	3.67***	4.60**	9.71***	0.50
	(7.92)	(6.61)	(2.41)	(3.43)	(0.91)
Total institutional holdings$_{t-1}$	−1.52***	−1.45***	−1.34***	−1.61***	0.17
	(−15.87)	(−11.71)	(−4.89)	(−8.81)	(1.60)
Number of institutions$_{t-1}$	−0.01***	−0.01***	−0.01***	−0.00	0.00
	(−4.91)	(−3.75)	(−2.47)	(−1.46)	(1.12)
Number of M&A deals$_{t-1}$	−0.05***	−0.06***	−0.01	−0.08***	−0.01
	(−6.30)	(−5.48)	(−0.29)	(−4.06)	(−0.95)
Number of issues$_{t-1}$	−0.61***	−0.63***	−0.35***	−0.77***	0.04
	(−11.82)	(−9.91)	(−2.77)	(−6.40)	(0.61)
Fixed effects	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	0.3747	0.3361	0.4375	0.4443	0.0028
Wald χ^2	11,236.34	6,540.85	1,506.84	3,396.72	66.17
Prob. > χ^2	0.0000	0.0000	0.0000	0.0000	0.3682
Number of observations	72,960	43,500	8,274	21,186	17,464

*, **, *** Indicate that coefficients are different from 0 at the two-tailed 10 percent, 5 percent, and 1 percent levels, respectively.

(continued on next page)

TABLE 4 (continued)

This table reports logistic regression models for the probability that a firm will lose analyst coverage in year *t*. In models 1–4, the dependent variable is a dummy variable equal to 1 for the 16,662 sample firm-year observations that lost analyst coverage in year *t*, and 0 for the universe of 87,516 covered firm-year observations. In model 5, for each sample firm, we select a covered firm matched on performance indicators and investor interest characteristics in year *t*–1. The propensity score matching method is a one-to-one nearest available neighbor matching within calipers, defined as 0.25 of the standard deviation of the propensity scores between the two groups. We find matches for 8,732 sample firm-year observations. Variables are winsorized at the 1 percent and 99 percent levels. Year, industry, and exchange fixed effects are included in all regression models. White's heteroscedasticity-adjusted Z-statistics are in parentheses.

for the sample and control firms. We cannot reject the null hypotheses of equality between medians for any of the variables at conventional significance levels (the lowest p-value is 0.19). We also run the logistic regression with the dependent variable equal to 1 when a firm is not covered in year *t*, or 0 when a firm is covered in year *t*, matched to a sample firm in year *t*−1. The results for model 5 of Table 4 show that none of the variables are significant and the pseudo-R^2 approximates zero, implying that our control firm selection procedures are effective. In the next section, we use this control group to examine the difference-in-differences effects of losing analyst coverage.

V. CONSEQUENCES OF LOSING ANALYST COVERAGE

After the firm loses all analyst coverage in year *t*, we compute the median differences in performance and investor interest between a sample firm and its control firm, over the five years following year *t*, from year *t*+1 to year *t*+5.[8] Since the trends in performance and investor interest are consistent across the five years, for conciseness, we report values for years *t*+1, *t*+3, and *t*+5 only.

Firm Performance and Investor Interest by Subperiods

Panel A of Table 5 reports differences in firm performance indicators and investor interest characteristics from year *t*+1 to year *t*+5, categorized by subperiods. Specifically, year *t* in which a firm loses analyst coverage is sorted into one of the three regulatory subperiods. The sample firms that lose coverage in the pre-Reg FD subperiod exhibit a significant deterioration in their performance indicators over the five years after losing coverage. Cash flow/total assets, total liabilities/total assets, and the composite performance index all consistently worsen relative to the control firms. These results suggest that analysts indeed have private information on the performance prospects of these firms and convey it to investors through their coverage decisions. In addition, investor interest characteristics, such as market capitalization, trading volume, bid-ask spread, institutional holdings, and number of institutions, significantly worsen relative to covered peers.[9] Given that the sample and control firms are matched on *ex ante* institutional presence, the drop in institutional holdings is likely to be driven by the loss in analyst coverage, consistent with O'Brien and Bhushan (1990).

---

[8] As described in Table 1, sample firm *i*-1994 is matched to a covered firm *j*-1994, based on information available in year 1993; sample firm *i*-1995 is matched to a covered firm *z*-1995, based on information available in 1994, and so on. The performance and investor interest indicators for observation *i*-1994 are observed in the five years following year *t*, in 1995 (year *t*+1), 1996 (year *t*+2), 1997 (year *t*+3), 1998 (year *t*+4), and 1999 (year *t*+5). This procedure is not only free from any look-ahead bias, but also avoids potential issues associated with overlapping periods since sample firms are re-matched to control firms every year, based on the information available up to that date.

[9] The median differences in number of M&A deals and number of issues between sample and control firms are insignificantly different from zero in all years at the conventional levels. For brevity, therefore, we do not tabulate the median differences for these variables.

TABLE 5

Differences in Performance Indicators and Investor Interest Characteristics after Losing Analyst Coverage Categorized by Subperiods and Number of Years with (without) Analyst Coverage

Panel A: Subperiods

	Pre-Reg FD, 1984–2000			Post-Reg FD, 2001–2002			Post-Reg AC, 2003–2008		
Year $t+n$	$t+1$	$t+3$	$t+5$	$t+1$	$t+3$	$t+5$	$t+1$	$t+3$	$t+5$
Performance Indicators									
Sales growth (%)	−0.24	0.62	−1.09	−2.70**	−2.13*	1.50	−1.82	−0.19	−2.14
	−1.50	*1.44*	*−1.32*	*−2.21*	*−1.81*	*0.42*	*−1.25*	*−0.38*	*−0.75*
Cash flow/Total assets (%)	−1.02***	−0.82***	−1.06***	−0.30	−0.24	0.79	−0.74*	0.28	−1.45
	−4.02	*−3.48*	*−2.50*	*−0.77*	*−0.65*	*1.44*	*−1.77*	*0.53*	*−0.75*
ROA (%)	−0.15*	−0.20	−0.03	−0.23	−0.04	1.07**	−0.52	0.22	−1.58
	−1.72	*−1.03*	*−0.65*	*−0.78*	*−0.42*	*2.00*	*−1.72*	*0.09*	*−1.10*
Total liabilities/Total assets	0.06***	0.07***	0.05***	0.00	−0.01	−0.01	0.00	−0.01	0.06
	25.66	*20.96*	*17.61*	*0.28*	*−0.23*	*−0.85*	*0.63*	*−1.31*	*1.13*
Ohlson's O-score	0.07	0.10*	0.17*	−0.20*	−0.20	−0.38*	0.05	−0.07	0.74
	1.52	*1.92*	*1.86*	*−1.77*	*−0.82*	*−1.82*	*1.00*	*−0.08*	*1.44*
Excess return (%)	3.64**	0.01	0.19	−0.28	−3.24	4.08*	2.13*	0.35	17.96
	2.39	*0.42*	*0.89*	*−0.13*	*−0.80*	*1.86*	*1.82*	*0.96*	*0.99*
Stock volatility (%)	0.00	−0.32	−0.35	−0.09	0.13	0.37	−0.23	0.41	−0.30
	0.50	*−0.62*	*−1.17*	*−0.03*	*0.17*	*1.03*	*−0.49*	*0.61*	*−0.78*
Performance index (%)	−1.24***	−1.45**	−2.62	3.91	4.98	13.75**	1.68	1.88	−9.86
	−2.93	*−2.09*	*−0.58*	*0.63*	*0.46*	*2.15*	*0.15*	*0.02*	*−0.41*
Investor Interest Characteristics									
Market capitalization	−3.53***	−4.67***	−6.41***	−8.38***	−21.25***	−12.75	−11.95***	−11.71**	−18.48**
	−5.47	*−4.36*	*−3.56*	*−3.97*	*−3.36*	*−1.29*	*−3.56*	*−2.24*	*−2.03*
B/M ratio	0.04***	0.02	0.03	0.02	0.03	−0.05	0.02	0.01	−0.15*
	3.06	*1.10*	*1.56*	*1.28*	*1.16*	*−1.08*	*1.21*	*0.39*	*−1.82*
Trading volume	−0.30***	−0.36***	−0.67***	−0.40***	−0.98***	−2.26***	−1.38***	−1.08***	−3.54***
	−7.30	*−6.55*	*−6.67*	*−3.98*	*−3.05*	*−3.36*	*−7.39*	*−4.17*	*−3.06*
Bid-ask spread (%)	0.53***	0.24***	0.27***	0.25***	0.13***	0.09***	0.23***	0.11***	0.25***
	8.73	*4.38*	*4.76*	*4.69*	*3.98*	*2.45*	*8.51*	*2.78*	*2.64*

(continued on next page)

TABLE 5 (continued)

Year $t+n$	Pre-Reg FD, 1984–2000			Post-Reg FD, 2001–2002			Post-Reg AC, 2003–2008		
	$t+1$	$t+3$	$t+5$	$t+1$	$t+3$	$t+5$	$t+1$	$t+3$	$t+5$
Total institutional holdings (%)	−2.32***	−2.69***	−3.73***	−3.02***	−5.08***	−3.58***	−3.38***	−3.05***	−4.58***
	−9.26	*−8.20*	*−7.55*	*−4.25*	*−4.58*	*−2.49*	*−5.46*	*−3.57*	*−2.49*
Number of institutions	−1***	−1***	−2***	−2***	−4***	−7***	−2***	−4***	−8***
	−9.85	*−7.82*	*−6.97*	*−5.64*	*−4.58*	*−3.34*	*−7.17*	*−4.66*	*−3.05*
Investor interest index	−0.12***	−0.15***	−0.19***	−0.12***	−0.11***	−0.17***	−0.11***	−0.11***	−0.51***
	−11.92	*−9.00*	*−7.21*	*−6.10*	*−3.42*	*−2.60*	*−6.93*	*−2.93*	*−3.02*

Panel B: Number of Years with Analyst Coverage

	Firms that Lose Analyst Coverage after Being Covered for					
	Three Years or Less			More Than Three Years		
Year $t+n$	$t+1$	$t+3$	$t+5$	$t+1$	$t+3$	$t+5$
Performance Indicators						
Sales growth (%)	−0.98	−0.36	−1.62	−2.12***	−1.21	0.34
	−0.17	*−0.19*	*−1.03*	*−3.13*	*−1.01*	*0.13*
Cash flow/Total assets (%)	−0.07	0.11	−1.08**	−0.49***	−0.19	−0.27
	−0.91	*0.99*	*−1.99*	*−3.06*	*−1.45*	*−0.89*
ROA (%)	−0.03	0.13	−0.36**	−0.43**	0.12	−0.12
	−0.46	*0.40*	*−1.95*	*−2.13*	*0.29*	*−0.02*
Total liabilities/Total assets	0.00	0.00	0.00	−0.01	−0.02	−0.01
	1.09	*0.82*	*0.46*	*−1.15*	*−1.68*	*−1.06*
Ohlson's O-score	−0.03	0.05	0.13*	−0.01	0.01	−0.01
	−0.52	*1.05*	*1.90*	*−0.85*	*0.56*	*−0.43*
Excess return (%)	1.82	1.29	−1.57	1.39	−0.30	1.87
	1.03	*0.24*	*−0.14*	*1.54*	*−0.95*	*1.53*
Stock volatility (%)	−0.06	−0.15	0.48**	−0.15	−0.14	−0.36
	−0.54	*−0.51*	*2.00*	*−0.01*	*−0.04*	*−0.96*
Performance index (%)	0.94	1.43	−2.45	0.96	1.64	4.78
	0.33	*0.25*	*−1.43*	*0.16*	*0.65*	*1.54*

(continued on next page)

TABLE 5 (continued)

	Firms that Lose Analyst Coverage after Being Covered for					
	Three Years or Less			More Than Three Years		
Year *t+n*	*t*+1	*t*+3	*t*+5	*t*+1	*t*+3	*t*+5
Investor Interest Characteristics						
Market capitalization	−3.56***	−4.40***	−5.30***	−5.36***	−3.36**	0.30
	−4.74	*−3.85*	*−3.68*	*−4.60*	*−2.47*	*0.74*
B/M ratio	0.03	−0.03	−0.05	0.04***	0.02	0.00
	1.49	*−1.37*	*−1.48*	*3.31*	*0.43*	*0.72*
Trading volume	−0.31***	−0.35***	−0.36***	−0.34***	−0.38***	−0.57***
	−7.90	*−4.78*	*−2.97*	*−5.76*	*−6.09*	*−4.12*
Bid-ask spread (%)	0.28***	0.15***	0.10***	0.32***	0.30***	0.32***
	8.80	*5.21*	*3.15*	*5.74*	*4.59*	*4.57*
Total institutional holdings (%)	−1.58***	−2.70***	−2.68***	−2.46***	−3.45***	−5.32***
	−6.54	*−7.39*	*−4.31*	*−8.06*	*−6.94*	*−7.65*
Number of institutions	−1***	−1***	−1***	−2***	−1***	−3***
	−5.84	*−6.25*	*−3.69*	*−9.77*	*−5.75*	*−6.30*
Investor interest index	−0.11***	−0.12***	−0.10***	−0.13***	−0.12***	−0.19***
	−11.37	*−7.51*	*−4.03*	*−9.89*	*−6.54*	*−5.93*

Panel C: Number of Years without Analyst Coverage

	Firms that Lose Analyst Coverage for					
	1st Year			2nd Year and Beyond		
Year *t+n*	*t*+1	*t*+3	*t*+5	*t*+1	*t*+3	*t*+5
Performance Indicators						
Sales growth (%)	−2.17**	−1.02	−2.24	−1.63	0.14	−0.16
	−2.07	*−0.81*	*−1.64*	*−1.39*	*0.36*	*−0.80*
Cash flow/Total assets (%)	−2.11***	−1.24***	−1.62***	−0.04	−0.08	0.37
	−6.35	*−3.99*	*−3.60*	*−0.32*	*−0.67*	*0.07*
ROA (%)	−0.76	−0.90	−0.73	0.18	0.01	0.36
	−0.40	*−0.70*	*−0.86*	*1.50*	*0.97*	*1.21*

(continued on next page)

TABLE 5 (continued)

	Firms that Lose Analyst Coverage for					
	1st Year			2nd Year and Beyond		
Year *t+n*	***t*+1**	***t*+3**	***t*+5**	***t*+1**	***t*+3**	***t*+5**
Total liabilities/Total assets	0.00	−0.02	−0.01	−0.02	−0.01	−0.01
	1.13	*−0.42*	*−0.32*	*−1.63*	*−1.53*	*−0.81*
Ohlson's O-score	0.31***	0.30***	0.39***	−0.11*	−0.02	0.05
	4.19	*3.43*	*3.37*	*−1.79*	*−0.11*	*0.70*
Excess return (%)	2.59**	−0.30	5.58	0.61	0.18	1.52
	1.93	*−0.23*	*1.58*	*1.14*	*0.81*	*0.52*
Stock volatility (%)	0.89***	0.64**	0.86***	−0.78***	−0.51***	−0.51*
	5.36	*2.17*	*3.45*	*−3.45*	*−2.66*	*−1.72*
Performance index (%)	−10.58***	−1.97**	−4.32**	3.58	2.74	1.60
	−5.16	*−2.10*	*−2.20*	*1.55*	*1.17*	*1.42*
Investor Interest Characteristics						
Market capitalization	−12.13***	−11.92***	−24.51***	−1.78***	−3.34***	−4.42***
	−9.06	*−5.29*	*−5.21*	*−3.11*	*−3.14*	*−2.76*
B/M ratio	0.03	−0.03	−0.03	0.04***	0.03*	0.00
	1.23	*−0.92*	*−1.21*	*3.66*	*1.72*	*0.05*
Trading volume	−0.24***	−0.22**	−0.03*	−0.49***	−0.61***	−0.60***
	−3.17	*−2.27*	*−1.73*	*−11.67*	*−9.03*	*−6.37*
Bid-ask spread (%)	0.52***	0.22***	0.19***	0.23***	0.27***	0.14***
	9.81	*4.61*	*4.03*	*6.76*	*6.63*	*3.13*
Total institutional holdings (%)	−3.97***	−4.29***	−4.61***	−1.22***	−2.89***	−4.32***
	−8.73	*−6.87*	*−5.96*	*−6.53*	*−7.96*	*−7.18*
Number of institutions	−2***	−3***	−3***	−1***	−1***	−2***
	−10.60	*−7.26*	*−5.65*	*−7.20*	*−7.60*	*−5.77*
Investor interest index	−0.15***	−0.14***	−0.18***	−0.09***	−0.12***	−0.16***
	−10.82	*−5.97*	*−4.49*	*−11.99*	*−9.52*	*−7.11*

*, **, *** Indicate that median values are different from zero at the two-tailed 10 percent, 5 percent, and 1 percent levels, respectively, according to the Wilcoxon matched-pairs signed-ranks test.

This table reports median differences in performance indicators and investor interest characteristics between sample firms and control firms, which are covered firms matched on performance indicators and investor interest characteristics in year $t-1$. Year t marks the calendar year when sample firms receive no analyst coverage. Z-statistics are shown in italic.

The pattern shifts dramatically after the regulatory changes. Starting from 2001, with the exception of sales growth, none of the performance indicators of sample firms are significantly different from their covered peers after the loss of analyst coverage. Sales growth significantly worsens for the sample firms (at the 5 percent level), but only in year *t*+1. The investor interest characteristics of sample firms continue to be significantly worse than those for covered peers.

Because it takes time for analyst coverage to add value to a firm, the negative effects of losing coverage might be greater for a firm that has been covered for a longer period. We therefore categorize sample firms depending on whether they had more than three years with analyst coverage before becoming orphans, where three years is the median duration of analyst coverage for the sample. In Panel B of Table 5, firms that lose coverage after having been covered for three years or less exhibit no significant subsequent declines in performance. Firms that have been covered longer than the median do exhibit a decline in sales growth, cash flow/total assets, and ROA in the first two years after the loss of coverage. The effects on performance are consistent with the hypothesis that the longer analysts choose to cover a firm, the more firm-specific information they generate. In contrast, the deterioration in investor interest characteristics is largely similar across both subsamples.

In Panel C of Table 5, we report performance indicators and investor interest characteristics by the number of years without analyst coverage. Performance indicators appear to be significantly worse than covered peers only for the first year without analyst coverage. In subsequent years, the performance is largely in line with their peers, the only exception being the difference in idiosyncratic volatility that is lower for sample firms. Again, investor interest characteristics are worse in all years following the loss of analyst coverage for both subsamples.

Overall, our results in Table 5 suggest that, before the changes in analyst regulation, analysts added value to firms by communicating private information to investors and by increasing investor recognition. The informational advantage was larger for firms that analysts had been covering for a longer period of time. After the regulatory changes, analysts' decisions to drop coverage on a firm are unlikely to convey private information about the firm's performance, but the orphaned firms continue to experience a significant decrease in their stock liquidity.

Firm Performance and Investor Interest by Analyst Characteristics

Table 6 analyzes whether the effects on a firm's performance and investor interest affected by the characteristics of the analyst(s) who decide to drop coverage on that firm. The analyst characteristics we examine are the number, type, and quality of the analyst(s) who last covered the firm in year *t*–1. Year *t* indicates the first year when a firm loses analyst coverage. As in Table 5, for conciseness, we report results only for alternate years.

Panel A of Table 6 categorizes sample firms by the number of analysts in year *t*–1. As shown in Table 2, sample firms typically experience a gradual, rather than precipitous, loss of analyst interest. From year *t*–1 to year *t*, 67 percent of sample firms lose the coverage of one last analyst. Only 20 percent lose the coverage of two analysts simultaneously, 7 percent of three analysts, and 6 percent of four or more analysts. Investor interest, but not performance, deteriorates when only one analyst drops coverage. From year *t*+1 to year *t*+5, market capitalization, trading volume, bid-ask spread, institutional holdings, and number of institutions all significantly worsen. The effects on investor interest are generally smaller when two analysts drop coverage than when only one drops coverage, suggesting that the larger the number of analysts who simultaneously drop coverage, the more likely the drop is related to performance-related factors. Consistent with this interpretation, after a firm loses the coverage of two or more analysts simultaneously, it exhibits a decline in several performance indicators. For example, sales growth, cash flow/total assets, and ROA for the sample firms are significantly lower than the control firms at the 1 percent level,

TABLE 6

Differences in Performance Indicators and Investor Interest Characteristics after Losing Analyst Coverage Categorized by Analyst Characteristics

Panel A: Number of Analysts Who Last Covered Sample Firms

	Firms that Lose the Coverage of 1 Analyst			Firms that Lose the Coverage of 2 or More Analysts		
Year *t+n*	***t+1***	***t+3***	***t+5***	***t+1***	***t+3***	***t+5***
Performance Indicators						
Sales growth (%)	−1.18	−0.64	−1.27	−3.65***	0.48	0.54
	−1.28	*−0.67*	*−1.11*	*−3.74*	*0.72*	*1.15*
Cash flow/Total assets (%)	−0.26	−0.57	−0.17**	−1.41***	−0.36**	−0.03
	−1.26	*−1.63*	*−2.05*	*−5.11*	*−2.57*	*−0.55*
ROA (%)	−0.04	−0.24	−0.38	−1.13***	−0.29	0.14
	−0.04	*−0.36*	*−1.24*	*−4.46*	*−1.78*	*0.39*
Total liabilities/Total assets	−0.01	−0.01	0.00	−0.03*	−0.03**	−0.04**
	−0.05	*−0.26*	*0.50*	*−1.85*	*−2.31*	*−2.27*
Ohlson's O-score	0.00	0.07*	0.22***	0.06	0.09	−0.04
	0.30	*1.78*	*3.03*	*1.12*	*0.96*	*−0.02*
Excess return (%)	−0.40	−0.23	2.77	0.54	0.63	2.40
	−0.66	*−0.12*	*1.21*	*1.67*	*1.27*	*0.60*
Stock volatility (%)	−0.70***	−0.43***	−0.24	1.01***	0.51**	0.63**
	−3.16	*−2.80*	*−0.90*	*5.00*	*2.39*	*2.43*
Performance index (%)	0.27	2.14	0.74	−2.17**	0.96	0.05
	0.42	*0.36*	*0.41*	*−2.06*	*0.91*	*0.43*
Investor Interest Characteristics						
Market capitalization	−4.76***	−6.49***	−10.00***	−5.31***	−3.66	−8.32
	−7.40	*−5.75*	*−5.28*	*−2.93*	*−1.53*	*−1.74*
B/M ratio	0.06***	0.02	0.01	−0.01	−0.01	−0.06
	4.45	*1.45*	*0.22*	*−0.04*	*−0.56*	*−1.42*
Trading volume	−0.54***	−0.67***	−0.64***	−0.15	−0.02	0.22
	−12.80	*−9.84*	*−7.21*	*−1.59*	*−1.14*	*0.30*
Bid-ask spread (%)	0.28***	0.28***	0.23***	0.36***	0.19***	0.04
	7.86	*6.50*	*5.04*	*8.32*	*4.78*	*1.03*
Total institutional holdings (%)	−1.69***	−3.88***	−5.29***	−2.92***	−2.36***	−2.47**
	−8.54	*−10.17*	*−9.51*	*−5.79*	*−3.62*	*−2.35*
Number of institutions	−1***	−2***	−3***	−1***	−2***	−1
	−10.86	*−9.69*	*−8.20*	*−5.46*	*−4.28*	*−2.15*
Investor interest index	−0.12***	−0.15***	−0.21***	−0.09***	−0.08***	−0.05*
	−14.83	*−11.29*	*−8.78*	*−6.75*	*−3.38*	*−1.89*

(continued on next page)

TABLE 6 (continued)

Panel B: Type of Analyst Who Last Covered Sample Firms

Year t+n	Firms that Lose Coverage of an Investment Bank Analyst t+1	t+3	t+5	Firms that Lose Coverage of Other Analyst t+1	t+3	t+5
Performance Indicators						
Sales growth (%)	−1.51***	−0.32	−0.84	−3.79*	−0.76	−0.19
	−2.62	*−0.12*	*−0.48*	*−1.85*	*−0.69*	*−0.36*
Cash flow/Total assets (%)	−0.62***	−0.47**	−0.18*	−0.48	−0.73	0.33
	−3.91	*−2.54*	*−1.93*	*−0.77*	*−1.27*	*0.64*
ROA (%)	−0.27**	−0.16	−0.27	−0.27	−0.64	0.17
	−2.53	*−0.88*	*−1.33*	*−0.35*	*−1.36*	*0.04*
Total liabilities/Total assets	−0.01	−0.02	−0.01	−0.03	0.00	0.00
	−0.62	*−1.23*	*−0.93*	*−1.62*	*0.97*	*0.08*
Ohlson's O-score	0.04	0.09*	0.14**	−0.03	0.01	0.02
	1.13	*1.88*	*2.27*	*−0.42*	*0.71*	*1.16*
Excess return (%)	1.64*	−0.33	3.31	0.04	1.05	−0.02
	1.68	*−0.33*	*1.35*	*1.25*	*1.28*	*−0.29*
Stock volatility (%)	−0.17	−0.27*	−0.01	−0.51	0.50*	0.18
	−0.08	*−1.78*	*−0.58*	*−0.32*	*1.77*	*0.06*
Performance index (%)	−0.64	0.35	0.82	0.96	−1.19	−2.88
	−1.27	*0.10*	*0.09*	*0.94*	*−0.38*	*−0.60*
Investor Interest Characteristics						
Market capitalization	−4.42***	−5.51***	−9.67***	−6.70***	−7.53***	−5.65**
	−6.45	*−4.77*	*−4.81*	*−4.80*	*−3.19*	*−2.40*
B/M ratio	0.04***	0.00	−0.01	0.04	0.04	0.04
	3.39	*0.39*	*−1.03*	*1.47*	*1.40*	*0.89*
Trading volume	−0.40***	−0.46***	−0.42***	−0.43***	−0.46***	−0.52***
	−10.47	*−8.24*	*−5.48*	*−4.52*	*−2.85*	*−2.83*
Bid-ask spread (%)	0.34***	0.26***	0.15***	0.22***	0.22***	0.32***
	10.32	*7.54*	*3.97*	*4.17*	*2.93*	*3.19*
Total institutional holdings (%)	−1.69***	−2.97***	−4.14***	−3.54***	−4.88***	−6.55***
	−8.51	*−9.06*	*−8.34*	*−6.43*	*−5.55*	*−4.18*
Number of institutions	−1***	−2***	−2***	−2***	−2***	−3***
	−10.92	*−9.66*	*−7.36*	*−5.11*	*−3.98*	*−3.22*
Investor interest index	−0.11***	−0.13***	−0.15***	−0.12***	−0.14***	−0.25***
	−14.51	*−10.32*	*−7.31*	*−6.88*	*−4.48*	*−4.44*

Panel C: Quality of the Analyst Who Last Covered Sample Firms

Year t+n	Firms that Lose Coverage of Inferior Analyst t+1	t+3	t+5	Firms that Lose Coverage of Superior Analyst t+1	t+3	t+5
Performance Indicators						
Sales growth (%)	−1.50	−0.93	−1.17	−2.01**	0.30	−0.06
	−1.47	*−0.98*	*−1.15*	*−2.09*	*0.73*	*−0.73*
Cash flow/Total assets (%)	−0.20**	−0.41	0.19	−0.92***	−0.62**	−0.65*
	−1.97	*−1.63*	*1.18*	*−3.62*	*−2.36*	*−1.90*
ROA (%)	−0.14	−0.26	−0.08	−0.41**	−0.24	−0.45
	−1.35	*−0.80*	*−1.15*	*−2.18*	*−1.01*	*−0.85*

(continued on next page)

TABLE 6 (continued)

Year *t+n*	Firms that Lose Coverage of Inferior Analyst			Firms that Lose Coverage of Superior Analyst		
	t+1	*t*+3	*t*+5	*t*+1	*t*+3	*t*+5
Total liabilities/Total assets	−0.01	−0.02	−0.01	−0.02	−0.02	−0.01
	−1.04	*−1.27*	*−0.87*	*−1.18*	*−0.90*	*−0.19*
Ohlson's O-score	−0.02	0.00	0.18**	0.07	0.16	0.09*
	−0.11	*1.25*	*2.06*	*1.35*	*1.62*	*1.79*
Excess return (%)	3.64***	0.55	3.35	−1.09	−0.63	2.10
	2.66	*1.20*	*0.90*	*−0.26*	*−0.02*	*0.92*
Stock volatility (%)	−0.18	−0.01	0.22	−0.35	−0.31	−0.08
	−0.52	*−0.46*	*1.10*	*−0.40*	*−0.96*	*−0.16*
Performance index (%)	0.51	3.12	1.15	−1.00	−0.30	−1.65
	0.28	*0.06*	*0.01*	*−1.43*	*−0.29*	*−0.49*
Investor Interest Characteristics						
Market capitalization	−4.81***	−5.53***	−7.60***	−4.94***	−5.80***	−11.27***
	−5.38	*−3.52*	*−3.46*	*−5.73*	*−4.27*	*−4.16*
B/M ratio	0.02	−0.02	−0.03	0.05***	0.03*	0.00
	1.28	*−0.55*	*−1.33*	*3.98*	*1.70*	*0.35*
Trading volume	−0.38***	−0.45***	−0.31***	−0.46***	−0.48***	−0.59***
	−7.28	*−5.39*	*−3.42*	*−9.07*	*−6.93*	*−5.14*
Bid-ask spread (%)	0.23***	0.23***	0.16***	0.39***	0.27***	0.17***
	7.02	*5.41*	*3.13*	*8.75*	*5.92*	*3.76*
Total institutional holdings (%)	−1.81***	−2.71***	−4.11***	−2.14***	−3.99***	−4.65***
	−7.58	*−6.43*	*−5.80*	*−7.16*	*−8.41*	*−7.28*
Number of institutions	−1***	−2***	−2***	−1***	−2***	−2***
	−8.01	*−7.01*	*−5.71*	*−9.09*	*−7.59*	*−5.55*
Investor interest index	−0.11***	−0.12***	−0.16***	−0.12***	−0.14***	−0.17***
	−10.21	*−6.91*	*−5.35*	*−12.61*	*−8.87*	*−6.46*

*, **, *** Indicate that median values are different from zero at the two-tailed 10 percent, 5 percent, and 1 percent levels, respectively, according to the Wilcoxon matched-pairs signed-ranks test.
This table reports median differences in performance indicators and investor interest characteristics between sample firms and control firms, categorized by analyst characteristics. Specifically, we look at the number (Panel A), type (Panel B), and quality (Panel C) of the analyst(s) who last covered sample firms in year *t*−1. Year *t* marks the first year when sample firms lose analyst coverage. Z-statistics are shown in italic.

implying that the presence of more analysts increases the depth of the information pool on a firm's future performance.[10]

Panel B of Table 6 categorizes sample firms by the type of the analyst who issued the last earnings estimate in year *t*−1. Specifically, we distinguish between investment bank analysts and other analysts who work for an independent broker or a paid-for firm. We examine paid-for analysts and independent broker analysts together, because a separate analysis of paid-for analysts would not be meaningful due to the modest number of observations. Clarke et al. (2011) document that

[10] When firms lose coverage of two or more analysts, the effects on investor interest appear to weaken over time. For some of these firms, analysts' decision to discontinue coverage may anticipate temporary performance issues. When these issues are resolved in later years, some firms may regain analyst coverage along with investor interest.

investment bank analysts are more likely to be perceived as influential by the market than analysts who work for independent brokers or paid-for research firms. There is some evidence that investment bank analysts drop firms because of unfavorable private information. In the years after the loss of analyst coverage, sales growth, cash flow/total assets, and ROA, all significantly decline relative to covered peers. When non-investment bank analysts drop coverage, only the differences in sales growth are significant. Interestingly, the effects on investor interest are largely similar across the two types of analysts.

Panel C of Table 6 categorizes sample firms by the quality of the analyst who issued the last earnings estimate in year *t*–1. As shown in Table 2, an Institutional Investor star analyst is rarely the last analyst to cover sample firms. To proxy for the quality of the last analyst, we use the analyst's forecasting ability as measured by PMAFE. This allows us to separate analysts into two almost equally populated groups of superior and inferior analysts. Specifically, a superior analyst has a negative PMAFE, indicating that her forecasting ability is better than the mean analyst ability in a year (Clement 1999). There is some evidence that superior analysts drop coverage owing to unfavorable private information. Sales growth, cash flow/total assets, and ROA for sample firms are significantly lower than their covered peers. Again, the effects on investor interest characteristics are largely similar across analyst quality. The similarity between Panels B and C is consistent with superior analysts being more likely to be affiliated with investment banks.

Firm Performance and Investor Interest by Firm's Information Environment

We expect that the loss of analyst coverage will exacerbate the investor interest effects for firms with an opaque information environment. We compute a reporting opacity index, as the first principal component of earnings volatility, absence of company earnings guidance, intangibles/total assets, auditor's qualified opinion, and absolute value and volatility of discretionary accruals over the last five years. Firms that report an above (below) median reporting opacity index are regarded as firms with an opaque (transparent) information environment.

Table 7 reports median differences in performance indicators and investor interest characteristics between the sample and control firms, categorized by the firm's information environment. All else equal, analyst coverage is more valuable for opaque firms. After the loss of analyst coverage, cash flow/total assets and Ohlson's O-score for opaque firms significantly worsen relative to their control firms (p-values are 0.02 and 0.05, respectively), while none of the differences in performance indicators are significant for transparent firms. Moreover, opaque firms experience a greater decline in investor interest than transparent firms. Our results complement Kirk (2011) who examines a sample of firms with no prior coverage and finds that firms with weaker information environments are more likely to buy paid-for coverage as they have the most to gain from analyst coverage.

Does the Loss of Analyst Coverage Predict Delisting?

An analysis of delisting dates indicates that, among the 4,836 sample firms that lost analyst coverage for the first time in year *t*, 3,374 (70 percent) subsequently delisted. About 24 percent of the sample firms delisted within two years from the initial loss of coverage, and 45 percent delisted within five years. Control firms report significantly lower percentages at the 1 percent level; 19 and 38 percent, respectively. We examine whether the loss of analyst coverage predicts the stock delisting in a multivariate framework, after we control for other factors. Specifically, we use a Cox proportional hazards model to compute the probability that a firm will be delisted after losing analyst coverage in year *t*. The observation units are the 8,732 sample firms and their control firms. All units enter the dataset in year *t* and are tracked over the next ten years until year *t*+10. Data are left-censored by construction, and some data can be also right-censored. Proportional hazards

TABLE 7

Differences in Performance Indicators and Investor Interest Characteristics after Losing Analyst Coverage Categorized by the Firm's Information Environment

	Firms with Transparent Information Environment			Firms with Opaque Information Environment		
Year *t+n*	***t+1***	***t+3***	***t+5***	***t+1***	***t+3***	***t+5***
Performance Indicators						
Sales growth (%)	−1.34	−0.69	−0.33	−1.34	−0.23	1.86
	−1.53	*−0.01*	*−0.46*	*−1.37*	*−0.36*	*1.09*
Cash flow/Total assets (%)	−0.23	−0.26	0.25	−0.98**	−0.44	0.58
	−1.08	*−1.17*	*0.13*	*−1.98*	*−1.53*	*0.23*
ROA (%)	0.24	0.25	0.32	−0.40	0.12	0.47
	1.04	*1.09*	*1.14*	*−0.70*	*0.17*	*1.00*
Total liabilities/Total assets	−0.02	−0.02	−0.03	−0.01	−0.03	−0.00
	−1.10	*−1.02*	*−1.18*	*−1.64*	*−1.10*	*−1.05*
Ohlson's O-score	−0.22	−0.27*	−0.15	0.12**	0.12*	0.00
	−1.60	*−1.70*	*−0.94*	*2.30*	*1.72*	*0.02*
Excess return (%)	0.21	0.98	0.54	2.24**	0.83	5.92**
	1.01	*1.27*	*0.54*	*2.09*	*1.25*	*2.40*
Stock volatility (%)	−0.48	−0.69	−0.34	−0.36	−0.46	0.49
	−0.98	*−1.60*	*−0.94*	*−0.60*	*−0.61*	*0.92*
Performance index (%)	1.03	1.30	0.26	−2.47	0.06	8.50*
	1.08	*1.15*	*0.09*	*−0.87*	*0.76*	*1.84*
Investor Interest Characteristics						
Market capitalization	−2.78	−3.80	−4.96	−8.09***	−9.01***	−6.67***
	−0.77	*−0.79*	*−1.14*	*−8.84*	*−6.36*	*−3.25*
B/M ratio	0.05***	0.03	0.02	0.03***	0.01	−0.03*
	3.11	*0.98*	*1.20*	*2.74*	*0.39*	*−1.78*
Trading volume	−0.27***	−0.29***	−0.40***	−0.70***	−0.88***	−0.86***
	−7.19	*−5.00*	*−3.62*	*−8.60*	*−6.42*	*−4.04*
Bid-ask spread (%)	0.10***	0.04	0.02	0.51***	0.38***	0.11**
	3.19	*0.92*	*0.25*	*11.29*	*7.39*	*2.50*
Total institutional holdings (%)	0.56	0.33	−0.46	−3.43***	−4.74***	−4.86***
	1.25	*−0.19*	*−0.60*	*−12.15*	*−10.00*	*−6.53*
Number of institutions	0***	−1**	−1*	−2***	−3***	−3***
	−3.02	*−2.43*	*−1.77*	*−13.49*	*−9.84*	*−5.26*
Investor interest index	−0.05***	−0.06***	−0.00	−0.17***	−0.20***	−0.17***
	−4.37	*−2.82*	*−0.71*	*−14.29*	*−9.78*	*−5.17*

*, **, *** Indicate that median values are different from zero at the two-tailed 10 percent, 5 percent, and 1 percent levels, respectively, according to the Wilcoxon matched-pairs signed-ranks test.

This table reports median differences in performance indicators and investor interest characteristics between sample firms and control firms, categorized by the firm's information environment. Firms with an opaque (transparent) information environment are firms with above (below) the median reporting opacity score, which is an index based on earnings volatility, absence of company earnings guidance, intangibles/total assets, auditor's qualified opinion, absolute value of discretionary accruals, and volatility of discretionary accruals over the last five years. In particular, earnings volatility is the standard deviation of EPS in the last five years. Absence of company earnings guidance is equal to 1 if a firm never provided earnings guidance over the last five years. Earnings guidance data come from First Call Historical Database. Intangibles/total assets is the average ratio between intangible assets and total assets over the last five years (Compustat items INTAN/AT). Auditor's qualified opinion is equal to 1 if the auditor issued at least one qualified opinion in the last five years (Compustat item AUOP = 2 or 5). Discretionary accruals are determined according to the industry-modified

(continued on next page)

TABLE 7 (continued)

Jones method (Jones 1991; Dechow et al. 1995). Industry is defined using the 48 Fama-French industries classified by four-digit SIC codes. Specifically, for each sample firm with an opaque (transparent) information environment, we select a control firm with an opaque (transparent) information environment that is still covered in year *t*. We find 7,223 control firms that match sample firms over performance indicators, investor interest characteristics, and information environment. Z-statistics are shown in italic.

models have, however, the methodological advantage of dealing with censoring issues by incorporating a positive probability that the event might never occur for cross-sectional units.

Our main variable of interest is a time-constant dummy, *SAMPLE FIRM*, which is equal to 1 for sample firms not covered by analysts in year *t*, or 0 for control firms covered in the same year. As a result of the survival design, the *SAMPLE FIRM* dummy indicates whether analysts' decision not to cover a firm in year *t* predicts a greater hazard of delisting for that firm than its control firm, after controlling for other publicly available factors. The other covariates include time-varying indicators of performance and investor interest, along with year, industry, and exchange fixed effects. In addition, we include a control variable for the continued coverage of sample firms, *COVERAGE* dummy, which is equal to 1 when a firm is covered in a given year *t+n*, or 0 otherwise. Loderer et al. (2009) document that the delisting hazard for a firm is a U-shaped function of its age. We thus include *TRADING AGE*, defined as the fraction of years from the first trading day, and its squared term.

Panel A of Table 8 reports coefficients and Lin and Wei's (1989) heteroscedasticity-robust Z-statistics from the Cox regressions. Model 1 relates to the entire sample period. Coefficients on the covariates have the expected signs. The likelihood that a firm will delist increases as the firm's financial health, its operating performance, and stock liquidity deteriorate. Even after controlling for other publicly available factors from year *t* to year *t*+10 however, the coefficient of the *SAMPLE FIRM* dummy is significant at the 1 percent level in predicting delisting. The results are robust to including the composite performance and investor interest indicators in place of the individual characteristics. Hazard ratios—computed as ($e^{coefficient} - 1$) but not reported in Table 8—suggest that analyst coverage is also economically important; the loss of analyst coverage for one year implies that a firm is 11 percent more likely to delist than its control firm over the next ten years. Models 2 to 4 categorize sample firms by subperiods. Losing analyst coverage in the pre-Reg FD subperiod significantly predicts delisting at the 1 percent level. In the next two subperiods, the coefficients of *SAMPLE FIRM* decrease in magnitude but remain statistically significant.

We run the Cox regression model 1 of Table 8, Panel A over various subsamples. For brevity, we report only the coefficient and Z-statistics for the *SAMPLE FIRM* dummy in Panels B and C of Table 8. Panel B of Table 8 examines the impact of losing analyst coverage on the delisting hazard, conditional on the number of years the sample firms spent with analyst coverage before the initial loss, and then without analyst coverage. A firm that loses analyst coverage after being covered for more than three years is 19 percent more likely to delist than its control firm. In addition, the longer a firm stays without coverage, the more likely it is to delist subsequently. In fact, the economic significance for the coefficient of the *SAMPLE FIRM* dummy increases as sample firms spend more years without analyst coverage. As a firm loses analyst coverage for one year, that firm is 8 percent more likely to delist than its control firm, while it is 17 percent more likely to delist as coverage is lost for four or more years, and 30 percent more likely to delist as coverage is lost for ten or more years. We note that these increases in probability are conditional on the firm choosing not to purchase coverage.

Panel C of Table 8 categorizes sample firms by the number, type, and quality of analyst(s) who last provided coverage in year *t*−1. The loss of analyst coverage in year *t* predicts a greater hazard of delisting for firms that lost the coverage of either two or more analysts, non-investment bank

TABLE 8

Cox Regression for the Probability that the Firm Will Be Delisted after Losing Analyst Coverage

Panel A: Hazard Likelihood of Delisting over Ten Years after Losing Coverage for One Year

	All Sample Period (1)	Pre-Reg FD, 1984–2000 (2)	Post-Reg FD, 2001–2002 (3)	Post-Reg AC, 2003–2008 (4)
SAMPLE FIRM dummy	0.10***	0.10***	0.04***	0.04***
	(15.58)	(12.55)	(2.67)	(2.59)
COVERAGE dummy	−0.02***	−0.03***	−0.02	−0.06***
	(−2.60)	(−3.43)	(−1.31)	(−3.63)
TRADING AGE	−0.07***	−0.07***	−0.01***	−0.01***
	(−67.84)	(−61.16)	(−4.22)	(−6.03)
TRADING AGE^2	0.00***	0.00***	0.00***	0.00***
	(47.55)	(47.37)	(3.05)	(4.70)
Performance Indicators				
Sales growth	−0.00	−0.01	−0.00	−0.01
	(−0.83)	(−0.83)	(−0.24)	(−1.27)
Cash flow/Total assets	−0.11**	−0.12*	0.02	−0.26**
	(−2.00)	(−1.85)	(0.23)	(−2.34)
ROA	−0.14***	−0.18***	−0.01	0.21**
	(−2.90)	(−3.19)	(−0.16)	(2.05)
Total liabilities/Total assets	0.02	0.07***	−0.05	−0.03
	(0.72)	(2.49)	(−1.21)	(−0.63)
Ohlson's O-score	0.02***	0.02***	0.01	0.00
	(5.29)	(5.14)	(1.04)	(0.43)
Excess return	−0.02***	−0.01	−0.03**	−0.02
	(−3.31)	(−1.38)	(−2.09)	(−0.94)
Stock volatility	0.08	0.09	0.14	0.31**
	(1.49)	(1.22)	(1.22)	(2.14)

(continued on next page)

American Accounting Association

TABLE 8 (continued)

	All Sample Period (1)	Pre-Reg FD, 1984–2000 (2)	Post-Reg FD, 2001–2002 (3)	Post-Reg AC, 2003–2008 (4)
Investor Interest Characteristics				
Ln(Market capitalization)	−0.02***	−0.02***	−0.02**	−0.01
	(−3.78)	(−2.89)	(−2.08)	(−0.75)
B/M ratio	0.00	0.01	0.00	0.00
	(1.02)	(0.87)	(0.07)	(0.19)
Ln(1 + Trading volume)	−0.04***	−0.03***	−0.02***	−0.04***
	(−11.38)	(−8.80)	(−3.71)	(−5.50)
Bid-ask spread	0.99***	0.92***	0.77**	0.73
	(8.91)	(6.93)	(1.94)	(1.32)
Total institutional holdings	−0.09***	−0.04*	−0.12***	−0.20***
	(−4.36)	(−1.65)	(−3.10)	(−4.56)
Number of institutions	−0.01**	−0.00	−0.04***	−0.02
	(−2.10)	(−0.60)	(−3.25)	(−1.16)
Number of M&A deals	−0.02***	−0.01***	−0.01	−0.00
	(−9.73)	(−7.67)	(−1.39)	(−1.02)
Number of issues	−0.00	−0.02*	−0.02	−0.02
	(−0.34)	(−1.64)	(−1.06)	(−0.89)
Fixed effects	Yes	Yes	Yes	Yes
Wald χ^2	20,586.91	37,761.94	63,755.70	5,877.63
Prob. > χ^2	0.0000	0.0000	0.0000	0.0000
Number of Observations	78,018	59,304	9,473	9,241

(continued on next page)

TABLE 8 (continued)

Panel B: Number of Years with (without) Analyst Coverage

	Number of Years with Analyst Coverage		Number of Years without Analyst Coverage				
	Three Years or Less	More than Three Years	1st Year	2nd Year	3rd Year	4th Year and Beyond	10th Year and Beyond
SAMPLE FIRM dummy	0.03***	0.18***	0.07***	0.05***	0.14***	0.16***	0.26***
	(2.92)	(19.93)	(6.46)	(3.58)	(7.18)	(12.93)	(7.85)
Other variables not reported							
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wald χ^2	11,426.45	14,059.09	9,245.72	5,034.89	3,801.81	7,478.18	2,273.16
Prob. > χ^2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Number of Observations	31,146	46,872	25,521	14,393	9,694	28,410	5,453

Panel C: Analyst Characteristics and Information Environment

	Number of Analysts Dropping Coverage		Type of the Analyst Dropping Coverage		Quality of the Analyst Dropping Coverage		Information Environment	
	1 Analyst	2 or More Analysts	Investment Bank	Other Firm	Inferior	Superior	Transparent	Opaque
SAMPLE FIRM dummy	0.09***	0.13***	0.09***	0.24***	0.09***	0.12***	0.07***	0.09***
	(11.13)	(11.00)	(12.15)	(13.33)	(9.23)	(13.18)	(6.21)	(9.51)
Other variables not reported								
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wald χ^2	13,833.42	13,480.60	18,094.74	8,778.11	10,729.52	10,200.05	11,942.35	20,808.85
Prob. > χ^2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
No. of Observations	54,151	23,867	66,707	11,311	39,412	38,329	23,891	36,381

*, **, *** Indicate that coefficients are different from zero at the two-tailed 10 percent, 5 percent, and 1 percent levels, respectively.

Panel A presents Cox regression results. The failure event is the stock's delisting. The duration analysis covers ten years from year *t* to year *t*+10. Year *t*, when sample firms receive no analyst coverage, is the baseline time. "Sample firm" is a time-invariant dummy variable equal to 1 for sample firms, and 0 for control firms matched on performance indicators and investor interest characteristics. All other covariates are time-varying variables. "Coverage" is a dummy variable equal to 1 if analysts issue at least one earnings forecast on firms during the calendar year *t*+*n*. "Trading age" is the number of years from the first trading day to the end of year *t*+*n*. Panels B and C present Cox regression results for model 1 in Panel A. Specifically, in Panel B, sample observations are categorized by the number of consecutive years with (without) analyst coverage. In Panel C, sample observations are categorized by analyst coverage characteristics. Firms with an opaque (transparent) information environment are sample firms with above (below) the median reporting opacity score. All variables are winsorized at the 1 percent and 99 percent levels. Year, industry, and exchange fixed effects are included in all regressions. Lin and Wei's (1989) heteroscedasticity-adjusted Z-statistics are in parentheses.

analysts, or superior analysts. Finally, we sort firms on their information environment. We find that the loss of coverage predicts a greater probability of delisting for stocks with opaque reporting.

VI. CONSEQUENCES OF REGAINING ANALYST COVERAGE

We next examine if the results found from examining the event of losing all analyst coverage hold as we analyze the complementary event of regaining analyst coverage. After losing analyst coverage for one year or more, 2,463 firms in our full sample regain analyst attention. It is reasonable to believe that the same factors that drive the analyst's decision to drop coverage on a firm can later explain her decision to resume coverage on that firm. Therefore, to predict the restoration of analyst coverage on a firm, we run model 1 of Table 4, in which now the dependent variable is equal to 1 when a sample firm that was not covered in year $t-1$ regains analyst coverage in year t, or 0 when a sample firm continues not to be covered in year t. Again, we use PSM over the 15 independent variables and fixed effects to select, for each firm that regains analyst coverage in year t, a control firm among the firms that continue to be uncovered. We identify 1,809 control firms.

Table 9 reports median differences in performance indicators and investor interest characteristics between sample firms that regain analyst coverage and their control firms, from year $t+1$ to year $t+5$. These results, categorized by subperiod, are largely consistent with those reported in Table 5. In the pre-Reg FD subperiod, after regaining analyst coverage, sample firms consistently outperform the control group in year $t+1$. Cash flow/total assets and ROA are significantly higher, while the total liabilities/total assets ratio is significantly lower than for the control firms, suggesting that analysts indeed have private information on the future prospects of these firms and convey it to investors through their decisions to resume coverage. In addition, the sample firms that regain coverage exhibit better market capitalizations, trading volumes, bid-ask spreads, and institutional presence than their peers. As in Panel A of Table 5, the pattern shifts considerably after the introduction of the new regulations. Few of the performance indicators are significantly different from their control firms after the reinstatement of analyst coverage. Investor interest proxies continue to improve significantly over time. Overall, these results are consistent with our earlier conclusion: after the regulatory changes, analysts continue to add value to a firm by enhancing its stock's recognition among investors.

VII. ROBUSTNESS CHECKS

In this section, we test the robustness of our results. Our results are qualitatively unchanged to including firm fixed effects in logistic regression models of Table 4. Analysts might base their coverage decisions on current factors, instead of on lagged indicators of performance and investor interest. We therefore repeat the PSM procedure by matching sample firms over the 15 independent variables computed in year t, instead of year $t-1$. Again, our results are qualitatively unchanged. In the remaining checks, we analyze: (1) partial and exogenous losses of analyst coverage; (2) a sample of firms that receive actual termination notices from analysts; and (3) the institutional investors who are likely to be affected in their investment decisions by the lack of analyst coverage.

Partial and Exogenous Losses of Analyst Coverage

What makes the loss of the last analyst(s) so important? We have no priors that the complete loss of the last two analysts, for example, signals more (or less) private information about the future performance of a firm than the partial loss of two analysts from five to three. However, to the extent that losing the last analyst(s) affects the investor recognition for a firm, we expect that the decrease in investor interest for a firm will be worse when it loses all analyst coverage and becomes an orphan than when it suffers only a partial reduction in analyst coverage.

TABLE 9

Differences in Performance Indicators and Investor Interest Characteristics after Regaining Analyst Coverage Categorized by Subperiods

	Pre-Reg FD Period 1984–2000			Post-Reg FD Period 2001–2002			Post-Reg AC Period 2003–2008		
Year $t+n$	$t+1$	$t+3$	$t+5$	$t+1$	$t+3$	$t+5$	$t+1$	$t+3$	$t+5$
Performance Indicators									
Sales growth (%)	−0.07	−2.36**	1.52	−1.18	−3.76	−0.44	2.63	1.34	0.91
	−0.14	*−2.38*	*0.04*	*−0.64*	*−1.13*	*−0.87*	*0.76*	*0.90*	*0.33*
Cash flow/Total assets (%)	2.93***	1.65***	−0.15	0.50	0.43	4.64**	0.92	−0.26	−0.78
	2.89	*2.59*	*−0.25*	*0.38*	*1.15*	*2.46*	*1.06*	*−1.36*	*−0.59*
ROA (%)	1.32***	0.06	0.63	1.37	1.31	2.32**	0.68	−0.33	−1.29
	2.77	*0.79*	*0.27*	*1.28*	*1.08*	*1.99*	*0.43*	*−1.46*	*−0.63*
Total liabilities/Total assets	−0.09***	−0.03	0.03	0.03	0.01	−0.04	0.01	−0.03*	0.02
	−4.05	*−0.97*	*0.73*	*1.22*	*0.83*	*−0.02*	*0.18*	*−1.84*	*0.10*
Ohlson's O-score	−0.10*	−0.39*	−0.03	−0.14	−0.41	−0.58*	−0.66**	−0.02	0.17
	−1.65	*−1.91*	*−0.61*	*−0.78*	*−1.60*	*−1.71*	*−2.43*	*−0.39*	*0.08*
Excess return (%)	−4.85	0.66	6.50	12.65	−3.08	−7.40	−5.16	−0.34	−4.08
	−1.26	*0.24*	*1.59*	*1.32*	*−0.65*	*−0.09*	*−1.58*	*−0.21*	*−0.66*
Stock volatility (%)	−0.53	0.06	−0.49	0.56	−0.85	−1.08	1.48**	0.31	0.59
	−1.57	*1.02*	*−0.90*	*0.00*	*−0.62*	*−0.91*	*2.12*	*1.15*	*0.37*
Performance index (%)	6.19**	5.51	0.98	2.90	1.29	19.66*	7.01	−8.88	−0.29
	2.47	*1.51*	*0.18*	*1.45*	*0.62*	*1.84*	*1.47*	*−0.61*	*−0.21*
Investor Interest Characteristics									
Market capitalization	5.37**	5.40	10.72	18.79***	30.98*	11.06	23.26***	20.43	31.12
	2.06	*1.56*	*1.35*	*2.60*	*1.67*	*0.53*	*2.60*	*1.38*	*1.24*
B/M ratio	−0.01	−0.02	0.01	−0.09*	0.06	0.10	−0.07***	−0.09**	−0.05
	−0.27	*−0.58*	*0.06*	*−1.66*	*0.15*	*1.02*	*−2.58*	*−2.38*	*−0.69*
Trading volume	0.38***	0.44**	0.33	1.55***	1.26	0.21	3.82***	6.04***	12.27***
	2.89	*2.45*	*1.58*	*3.20*	*1.37*	*0.36*	*5.53*	*3.73*	*3.84*

(continued on next page)

TABLE 9 (continued)

	Pre-Reg FD Period 1984–2000			Post-Reg FD Period 2001–2002			Post-Reg AC Period 2003–2008		
Year $t+n$	$t+1$	$t+3$	$t+5$	$t+1$	$t+3$	$t+5$	$t+1$	$t+3$	$t+5$
Bid-ask spread (%)	−0.92***	−0.43***	−0.26***	−0.52***	−0.26***	−0.04	−0.42***	−0.28***	−0.30**
	−7.87	*−4.47*	*−2.71*	*−3.98*	*−3.78*	*−0.80*	*−7.65*	*−3.94*	*−2.00*
Total institutional holdings (%)	3.32***	4.23***	5.24***	7.88***	7.73**	7.54	6.53***	2.79**	12.18
	5.27	*4.40*	*4.63*	*4.59*	*2.52*	*0.96*	*4.60*	*2.17*	*1.62*
Number of institutions	2***	3***	4***	4***	5**	4	8***	5***	12***
	5.24	*4.67*	*4.36*	*3.91*	*2.47*	*1.50*	*6.34*	*3.15*	*2.56*
Investor interest index	0.20***	0.15***	0.19***	0.25***	0.24**	0.18	0.32***	0.36***	0.35
	7.78	*4.56*	*4.14*	*5.19*	*2.46*	*0.62*	*5.67*	*3.07*	*1.76*

*, **, *** Indicate that median values are different from zero at the two-tailed 10 percent, 5 percent, and 1 percent levels, respectively, according to the Wilcoxon matched-pairs signed-ranks test.

This table reports median differences in performance indicators and investor interest characteristics between firms that regain analyst coverage and control firms that are uncovered firms matched on performance indicators and investor interest characteristics in year $t-1$. Year t marks the calendar year when analyst coverage is resumed. Z-statistics are shown in italic.

To assess the importance of the last analyst(s), we perform a one-to-one matching of uncovered firms to covered firms on the decrease in analyst following, from year $t-1$ to year t. Specifically, for each sample firm that completely loses coverage by Δn analysts in year t, we select a control firm that loses coverage by the same number Δn of analysts in year t but still remains covered by at least one analyst. Again, we run model 1 of Table 4 over the 15 independent variables and fixed effects to estimate propensity scores, and use PSM to match firms on Δn. We find 2,639 control firms. The first set of columns of Table 10 reports the median differences between sample firms that experience a complete loss of analyst coverage and control firms that experience a partial loss in year t. In the five years after a complete loss of coverage, only stock volatility significantly worsens relative to a partial loss, indicating that there is only a modest amount of incremental private information when the firm completely loses all coverage than when it remains covered by at least one analyst. As expected, orphaned firms experience a significantly greater deterioration in investor interest proxies than firms that only partly lose coverage. Consistent with the investor recognition story, this result adds to the literature on the partial losses of analyst coverage.

To make sure that our research design corrects for possible endogeneity issues, we repeat the analysis on a subsample of firms that lose all analyst coverage for exogenous reasons. Similarly to Hong and Kacperczyk (2010), we search for events when the research department of the only brokerage house covering a firm closes due to concentration or corporate restructuring, making the firm an orphan in year t. We find 270 firms that are orphaned for exogenous reasons. Most events are confined either to small brokerage firms or geography-specialized divisions of larger brokerage houses. After manually checking that these events indeed correspond to exogenous losses of coverage, we follow the 270 firms over time leading to a subsample of 943 firm-year observations. For each subsample firm that remains uncovered in year t due to exogenous reasons, we select a firm that is covered in year t. Using PSM again, we find 665 control firms that are covered firms with comparable performance and investor interest characteristics. The second set of columns of Table 10 reports the median differences in the five years following year t. The results confirm that the event of losing analyst coverage for subsample firms is indeed exogenous and not performance-related. No performance indicator consistently deteriorates in the five-year period following the loss. Consistent with the investor recognition story, however, some proxies for stock liquidity worsen: bid-ask spread widens in year $t+1$, while total institutional holdings and the number of institutions drop dramatically relative to control firms, suggesting that the lack of analyst coverage drives institutions to sell their holdings.

Termination Notices

We analyze a large sample of termination notices issued in compliance with NASD Rule 2711. As mentioned earlier, paragraph (f)(5) of this rule requires an analyst who intends to drop the coverage of a firm to issue a final report specifying the rationale for the termination.[11] From Thomson Research Investext database, which stores the PDF files of the original analyst reports, we manually collect data on 7,038 termination notices issued between 2003 and 2008. About 44 percent of the termination notices refer to the analyst departure, 29 percent to redirection of research

[11] According to the text of Rule 2711 paragraph (f)(5): "If a member intends to terminate its research coverage of a subject company, notice of this termination must be made. The member must make available a final research report on the subject company using the means of dissemination equivalent to those it ordinarily uses to provide the customer with its research reports on the subject company. The report must be comparable in scope and detail to prior research reports and must include a final recommendation or rating, unless it is impracticable for the member to produce a comparable report (e.g., if the research analyst covering the subject company or sector has left the member or if the member terminates coverage of the industry or sector). If it is impracticable to produce a final recommendation or rating, the final research report must disclose the member's rationale for the decision to terminate coverage."

TABLE 10

Differences in Performance Indicators and Investor Interest Characteristics for Subsamples of Partial and Exogenous Losses of Analyst Coverage

	Partial Losses of Analyst Coverage			Exogenous Losses of Analyst Coverage		
Year $t+n$	$t+1$	$t+3$	$t+5$	$t+1$	$t+3$	$t+5$
Performance Indicators						
Sales growth (%)	0.32	1.47	−0.40	−5.54	−0.69	−14.05
	0.17	*0.85*	*−0.23*	*−1.47*	*−0.20*	*−0.92*
Cash flow/total assets (%)	0.04	−0.62	−0.19	−1.04*	−1.69*	0.17
	0.03	*−0.85*	*−0.98*	*−1.78*	*−1.77*	*−0.12*
ROA (%)	0.21	0.15	−0.13	−1.09	−0.21	−0.55
	0.92	*0.52*	*−0.34*	*−1.06*	*−1.08*	*−0.10*
Total liabilities/total assets	−0.01	−0.01	−0.01	0.03	0.02	0.10
	−0.05	*−0.34*	*−0.18*	*0.72*	*0.70*	*1.52*
Ohlson's O-score	−0.01	0.00	−0.06	0.51*	1.23**	0.76
	−0.28	*0.47*	*−0.74*	*1.87*	*2.07*	*1.03*
Excess return (%)	−1.13	−1.48	0.27	−5.41	1.83	11.10
	−0.20	*−0.65*	*0.02*	*−0.22*	*0.12*	*1.59*
Stock volatility (%)	1.14***	0.61	0.35*	0.70*	0.45	0.94
	3.73	*1.42*	*1.81*	*1.88*	*0.51*	*0.50*
Performance index (%)	2.36	−0.20	1.46	−1.36	−1.68*	−0.38
	0.60	*−0.44*	*0.49*	*−1.60*	*−1.92*	*−0.44*
Investor Interest Characteristics						
Market capitalization	−0.58	−0.71	−4.65**	−12.01	−16.20	−0.09
	−1.27	*−1.21*	*−2.09*	*−1.60*	*−1.41*	*−0.04*
B/M ratio	0.01	0.01	0.01	−0.17*	0.00	−0.07
	0.73	*0.61*	*0.30*	*−1.90*	*0.70*	*−0.87*
Trading volume	−0.72***	−0.87***	−0.48***	−0.03	0.08	−0.58
	−7.79	*−5.91*	*−3.04*	*−0.55*	*0.18*	*−0.29*
Bid-ask spread (%)	0.16***	0.12**	0.08*	0.58***	0.17	−0.03
	2.31	*2.09*	*1.65*	*2.66*	*0.39*	*−0.84*
Total institutional holdings (%)	−3.41***	−3.55***	−4.30***	−6.09***	−8.18***	−2.24***
	−7.26	*−5.89*	*−4.63*	*−3.81*	*−3.67*	*−3.06*
Number of institutions	−1.5***	−1***	−3***	−2.5***	−3.5***	0
	−7.64	*−5.21*	*−3.81*	*−3.89*	*−3.52*	*0.19*
Investor interest index	−0.10***	−0.13***	−0.14***	−0.19***	−0.00	−0.18
	−7.85	*−5.49*	*−3.72*	*−3.08*	*−1.58*	*−0.77*

*, **, *** Indicate that median values are different from zero at the two-tailed 10 percent, 5 percent, and 1 percent levels, respectively, according to the Wilcoxon matched-pairs signed-ranks test.

The first set of columns reports the median differences between sample firms that experience a complete loss of analyst coverage and control firms that experience a partial loss in year t. Specifically, for each sample firm that completely loses the coverage from Δn analysts in year t, we select a control firm that loses the coverage from the same number Δn of analysts in year t but still remains covered by at least one. We find 2,639 control firms that match sample firms over performance indicators, investor interest characteristics, and the change in analyst following, Δn. The second set of columns reports the median differences between a subsample of firms that lose all analyst coverage in year t due to exogenous reasons and control firms that are covered firms in year t matched on performance indicators and investor interest characteristics. A firm is defined as having lost all analyst coverage due to exogenous reasons when the research department of the only brokerage house covering the firm closes due to concentration or corporate restructuring. For our subsample of 943 firms experiencing exogenous losses, we find 665 control firms. Z-statistics are shown in italic.

efforts, 25 percent to stock delisting, and the remaining 2 percent of the notices disclose no reason.[12] There are 399 cases of analysts' departures that result in complete, exogenous losses of coverage. We analyze the changes in stock liquidity for this subsample over a ±365 day window surrounding the termination notice day. We find that the median trading volume declines by −3.15 percent, while the median bid-ask spread significantly widens in the 365-day period following the termination notice. Also, the institutional presence in the median firm decreases significantly with total holdings, dropping by −5.49 percent and the number of institutions reducing by 3.

Analyst Coverage and Institutional Presence

Institutional presence in a firm declines after the firm becomes an orphan either for endogenous or exogenous reasons. Which institutions are likely to follow the analysts? Institutions are not all equal: characteristics, such as size and industry, can make some institutions more sensitive to the loss of analyst coverage than others. Some types of money managers may use stock visibility, proxied by analyst coverage, as a factor in their portfolio allocation decisions.[13] We examine which institutions typically depend on sell-side analyst research at the end of 2008.

We categorize institutions by type and portfolio size quintiles.[14] In untabulated results, large institutions in the highest size quintile appear to rely more on their own research departments than small institutions in the lowest size quintile. The median stock held by large institutions receives a total number of 45 reports per year from 11 analysts, while the median stock held by small institutions receives more reports (62) from a higher number of analysts (16). Large institutions also hold more stocks that have no analyst coverage than the small institutional portfolios. Across the types of institution, large banks invest the most in uncovered stocks (8.68 percent), while small pension funds invest the least in these stocks (2.01 percent). The fact that small investment advisors hold a significant proportion of uncovered stocks (7.46 percent) is not surprising, given that the "Investment advisors" category is quite heterogeneous and includes hedge funds and private equity funds that might strategically invest in uncovered and under-covered firms. Interestingly, small institutions mostly invest in stocks covered by star analysts. Overall, the analysts' decision to stop covering a stock is likely to affect the investment in that stock by small institutions, such as small pension funds, mutual funds, banks, and insurance companies.

VIII. CONCLUSIONS

The popular press has highlighted how firms, especially small- and mid-capitalization firms are increasingly losing analyst coverage (*The Economist* 2009). In 2005, the NASDAQ and Reuters

---

[12] The reason for terminating coverage of a given firm while redirecting research efforts elsewhere is likely to be related to the firm's performance and/or investor interest. However, analysts appear reluctant to explicitly provide the firm-specific reasons behind their decisions to terminate coverage. They use expressions like "the stock no longer fits the coverage universe" and, only in less than 2 percent of notices, do they specify why, in their view, the stock is unfit for coverage. In general, the final reports do not appear to be comparable to prior research reports, either in scope or detail. Most of them revise the next-to-last recommendation to "Not rated."

[13] We are unaware of any regulatory or internal provision requiring an institution to invest only in covered stocks, although we cannot entirely eliminate the possibility.

[14] From Exchange Act Form 13f holdings filings, we remove those securities that cannot be identified by a PERMNO on the CRSP database. Securities without PERMNOs may be listed in foreign markets and therefore not covered by U.S. analysts. We also remove exchange-traded fund shares and mutual fund shares, which are generally not covered by analysts. To identify the type of institution, we use the *TYPECODE* variable in the 13f filings: Banks and trusts = 1; Insurance companies = 2; Mutual funds = 3; Investment advisors = 4; All others = 5. The "All others" category includes pension funds, endowments, and foundations. Thomson acknowledges serious classification errors starting from 1998, when many of the institutions were improperly classified as "All others." Hence, we manually check the institutions originally classified as "All others" and distribute them in one of the five categories.

created the Independent Research Network (IRN) to help public companies to obtain research. NASDAQ and Reuters motivated the launch of the IRN by citing the fact that about 35 percent of all public firms listed in the U.S. markets lack analyst coverage, and that 691 firms representing over 17 percent of the entire universe of covered firms lost all analyst coverage between 2002 and 2005.[15] At the same time, paid-for research boutiques expanded their business by offering coverage to firms neglected by sell-side analysts in return for fees.

This paper analyzes firm performance and investor interest after losing all analyst coverage to test if analysts add value to a firm by signaling private information on the firm's performance and/or by increasing its visibility to investors. We find evidence of both in a sample of 16,662 firm-year observations without any analyst coverage for at least one year, from 1984 to 2008. In addition, we find that after the introduction of regulations that eroded analyst informational advantages, sample firms no longer experience a significant difference in performance indicators relative to covered peers. However, they continue to experience a significant deterioration of bid-ask spreads, trading volumes, and institutional presence relative to their peers. Moreover, sample firms that lose all analyst coverage for one year are significantly more likely to delist than their peers.

Our results have important policy implications. By eliminating investment-banking-derived compensation for analysts, the new analyst regulations changed the economic incentives in the research industry. Analysts employed in traditional soft-dollar brokerage houses now may have limited incentives to cover small- and mid-cap firms since these firms are unlikely to generate sufficient order flows to justify the costs of maintaining coverage. Our results offer a rationale for small- and mid-cap firms seeking hard-dollar research. As the paid-for research industry evolves, we may see a debate surrounding the appropriate model for equity research similar to the one surrounding credit-rating agencies. One policy implication of our work hinges on how research efforts and compensation should be structured in terms of quality, independence, and objectivity.

REFERENCES

Altınkılıç, O., and R. S. Hansen. 2009. On the information role of stock recommendation revisions. *Journal of Accounting and Economics* 48: 17–36.

Amihud, Y., and H. Mendelson. 1986. Asset pricing and the bid-ask spread. *Journal of Financial Economics* 17: 223–249.

Baker, H. K., J. R. Nofsinger, and D. G. Weaver. 2002. International cross listing and visibility. *Journal of Financial and Quantitative Analysis* 37: 495–521.

Barber, B. M., R. Lehavy, M. McNichols, and B. Trueman. 2006. Buys, holds, and sells: The distribution of investment banks' stock ratings and the implications for the profitability of analysts' recommendations. *Journal of Accounting and Economics* 41: 87–117.

Bhushan, R. 1989. Firm characteristics and analyst following. *Journal of Accounting and Economics* 11: 255–274.

Bradshaw, M. T. 2009. Analyst information processing, financial regulation, and academic research. *The Accounting Review* 84: 1073–1083.

Brennan, M. J., and A. Subrahmanyam. 1995. Investment analysis and price formation in securities markets. *Journal of Financial Economics* 38: 361–381.

Bushee, B. J., and G. S. Miller. 2010. *Investor Relations, Firm Visibility, and Investor Following*. Working paper, University of Pennsylvania and University of Michigan.

[15] See the NASDAQ press release of June 7, 2005. IRN operated as a research matchmaker, connecting firms that seek coverage for their stocks with independent research firms and passing fees from the firms to the research firms. The annual charge for the firms was $100,000 for three years, with the three-year commitment designed to keep the firms from pulling out in case of unfavorable research.

Chen, C.-Y., and P. F. Chen. 2009. NASD rule 2711 and changes in analysts' independence in making stock recommendations. *The Accounting Review* 84: 1041–1071.
Clarke, J., A. Khorana, A. Patel, and P. R. Rau. 2007. The impact of all-star analyst job changes on their coverage choices and investment banking deal flow. *Journal of Financial Economics* 84: 713–737.
Clarke, J., A. Khorana, A. Patel, and P. R. Rau. 2011. Independents' day? Analyst behavior surrounding the Global Settlement. *Annals of Finance* 7: 529–547.
Clement, M. B. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27: 285–303.
Cliff, M. T., and D. J. Denis. 2004. Do IPO firms purchase analyst coverage with underpricing? *Journal of Finance* 59: 2871–2901.
Collins, D. W., S. P. Kothari, and J. Dawson Rayburn. 1987. Firm size and the information content of prices with respect to earnings. *Journal of Accounting and Economics* 9: 111–138.
Cooper, M. J., H. Gulen, and P. R. Rau. 2005. Changing names with style: Mutual fund name changes and their effects on fund flows. *Journal of Finance* 60: 2825–2858.
Dechow, P., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70: 193–225.
Easley, D., M. O'Hara, and J. Paperman. 1998. Financial analysts and information-based trade. *Journal of Financial Markets* 1: 175–201.
Economist, The. 2009. High-speed slide. (November 14): 85–26.
Ellul, A., and M. Panayides. 2009. *Do Financial Analysts Restrain Insiders' Informational Advantage?* Working paper. Indiana University and The University of Utah.
Frankel, R., S. P. Kothari, and J. Weber. 2006. Determinants of the informativeness of analyst research. *Journal of Accounting and Economics* 41: 29–54.
Gao, X., and J. R. Ritter. 2010. The marketing of seasoned equity offerings. *Journal of Financial Economics* 97: 33–52.
Gintschel, A., and S. Markov. 2004. The effectiveness of Regulation FD. *Journal of Accounting and Economics* 37: 293–314.
Hong, H., and M. Kacperczyk. 2010. Competition and bias. *Quarterly Journal of Economics* 125: 1683–1725.
Irvine, P. J. 2001. Do analysts generate trade for their firms? Evidence from the Toronto stock exchange. *Journal of Accounting and Economics* 30: 209–226.
Irvine, P. J. 2003. The incremental impact of analyst initiation of coverage. *Journal of Corporate Finance* 9: 431–451.
Irvine, P. J. 2004. Analysts' forecasts and brokerage-firm trading. *The Accounting Review* 79: 125–149.
James, C., and J. Karceski. 2006. Strength of analyst coverage following IPOs. *Journal of Financial Economics* 82: 1–34.
Jegadeesh, N., J. Kim, S. D. Krische, and C. M. C. Lee. 2004. Analyzing the analysts: When do recommendations add value? *Journal of Finance* 59: 1083–1124.
Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29: 193–228.
Kecskés, A., and K. L. Womack. 2007. *Changes in Analyst Coverage: Does the Stock Market Overreact?* Working paper, University of Toronto and Dartmouth College.
Kelly, B., and A. Ljungqvist. 2011. Testing asymmetric-information asset pricing models. *Review of Financial Studies* 25: 1366–1413.
Kirk, M. 2011. Research for sale: Determinants and consequences of paid-for analyst research. *Journal of Financial Economics* 100: 182–200.
Krigman, L., W. H. Shaw, and K. L. Womack. 2001. Why do firms switch underwriters? *Journal of Financial Economics* 60: 245–284.
Lehavy, R., and R. Sloan. 2008. Investor recognition and stock returns. *Review of Accounting Studies* 13: 327–361.
Lehavy, R., F. Li, and K. Merkley. 2011. The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review* 86: 1087–1115.

Lin, D., and L. Wei. 1989. Robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association* 84: 1074–1078.

Loderer, C., K. Neusser, and U. Waelchli. 2009. *Firm Age and Survival*. Working paper, University of Bern.

Loughran, T., and J. Ritter. 2004. Why has IPO underpricing changed over time? *Financial Management* 33: 5–37.

McNichols, M. F., and P. O'Brien. 1997. Self-selection and analyst coverage. *Journal of Accounting Research* 35: 167–199.

Mehran, H., and S. Peristiani. 2010. Financial visibility and the decision to go private. *Review of Financial Studies* 23: 519–547.

Merton, R. C. 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42: 483–510.

O'Brien, P., and R. Bhushan. 1990. Analyst following and institutional ownership. *Journal of Accounting Research* 28 (Supplement): 55–76.

Ohlson, J. A. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18: 109–131.

Pietroski, J. D., and D. Roulstone. 2004. The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices. *The Accounting Review* 79: 1119–1151.

Rajan, R., and H. Servaes. 1997. Analyst following of initial public offerings. *Journal of Finance* 52: 507–529.

Rosenbaum, P. R., and D. B. Rubin. 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician* 39: 33–38.

Shumway, T. 2001. Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business* 74: 101–124.

Villalonga, B. 2004. Does diversification cause the "diversification discount"? *Financial Management* 33: 5–27.

Weild, D., and E. Kim. 2009. A wake-up call for America. *Grant Thornton Capital Markets Series* (November).

Zmijewski, M. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* 22: 59–82.

APPENDIX A

Variable	Definition
Sales growth	Change in sales (Compustat item SALE) from the prior year.
Cash flow/total assets	Operating cash flow divided by total assets [(IB + DP)/AT].
ROA	Net income divided by total assets (NI/AT).
Total liabilities/Total assets	Total liabilities divided by total assets (LT/AT).
Ohlson's O-score	Measure of financial distress, defined as $-1.32 - 0.407 \times \log(\text{total assets}) + 6.03 \times (\text{total liabilities/total assets}) - 1.43 \times (\text{working capital/total assets}) + 0.076 \times (\text{current liabilities/current assets}) - 1.72 \times (1 \text{ if total liabilities} > \text{total assets, else } 0) - 2.37 \times (\text{net income/total assets}) - 1.83 \times (\text{funds from operations/total liabilities}) + 0.285 \times (1 \text{ if net loss for last two years, else } 0) - 0.521 \times (\text{change in net income/sum of absolute values of net income})$. That is, $-1.32 - 0.407 \times \log(\text{AT in constant 2005 dollars}) + 6.03 \times (\text{LT/AT}) - 1.43 \times (\text{WCAP/AT}) + 0.076 \times (\text{LCT/ACT}) - 1.72 \times (1 \text{ if LT} > \text{AT, else } 0) - 2.37 \times (\text{NI/AT}) - 1.83 \times (\text{OANCF/LT}) + 0.285 \times (1 \text{ if NI} < 0 \text{ for last two years, else } 0) - 0.521 \times (\text{change in NI/sum of absolute values of NI})$.
Excess return	Cumulative stock monthly return, adjusted for the CRSP NYSE/AMEX/NASDAQ value-weighted index.
Stock volatility	Idiosyncratic standard deviation of stock monthly returns.
Market capitalization	Common shares outstanding multiplied by fiscal year closing price (CSHO × PRCC_F), expressed in $ millions.
B/M ratio	Ratio (common equity + deferred taxes + investment tax credit − preferred stock)/market capitalization, or (CEQ + TXDB + ITCB-PSTK)/(CSHO × PRCC_F).
Trading volume	Total number of shares traded in the year. We use the algorithm in Gao and Ritter (2010) to adjust the trading volume of NASDAQ-listed stocks and avoid double-counting NASDAQ volume relative to NYSE/AMEX volume.
Bid-ask spread	Annual average of daily differences between the closing bid and ask prices scaled by the mid-range closing price.
Institutional holdings and Number of institutions	Total holdings and total number of institutional investors reporting in the fourth quarter, respectively. Data on institutional holdings come from the 13f Institutional Holdings database.
Number of M&A deals	Cumulative number of M&A deals in the last three years, respectively.
Number of issues	Cumulative number of new equity issues in the last three years. Data on M&A deals and new equity issues come from the SDC New Issues database.

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 707–737

American Accounting Association
DOI: 10.2308/accr-50314

Group Audits, Group-Level Controls, and Component Materiality: How Much Auditing Is Enough?

Trevor R. Stewart
Deloitte (Retired); Rutgers, The State University of New Jersey; and VU University Amsterdam

William R. Kinney, Jr.
The University of Texas at Austin

ABSTRACT: Auditing standards now mandate that group auditors determine and implement appropriate component materiality amounts, which ultimately affect group audit scope, reliability, and value. However, standards are silent about how these amounts should be determined and methods being used in practice vary widely, lack theoretical support, and may either fail to meet the audit objective or do so at excessive cost. We develop a Bayesian group audit model that generalizes and extends the single-component *audit risk model* to aggregate assurance across multiple components. The model formally incorporates group auditor knowledge of group-level structure, controls, and context as well as component-level constraints imposed by statutory audit or other requirements. Application of the model yields component materiality amounts that achieve the group auditor's overall assurance objective by finding the optimal solution on an *efficient materiality frontier*. Numerical results suggest group-level controls and structured subgroups of components are central to efficient group audits.

Keywords: *aggregation; audit assurance; Bayes' rule; component materiality; group audits; group-level controls; ISA 600.*

Data Availability: *Upon request, Dr. Stewart will provide Excel-based software that facilitates exploration and application of the model described in this paper.*

We gratefully acknowledge the advice and help we have received from many individuals. We especially thank Hans Moors for his invaluable technical guidance, Leo Strijbosch for designing and running hundreds of simulations to help validate key results, Paul Dunmore for helping to resolve certain computational problems, Roman Chychyla for his mathematical advice, Matt Ege and Marcy Shepardson for research assistance, and Paul van Batenburg for his support throughout. We also gratefully acknowledge the helpful comments on earlier drafts of the paper by Andy Bailey, Wolf Böhm, Ed Broeze, Dave Burgstahler, Judson Caskey, Bill Felix, Lucas Hoogduin, Al Leitch, Zoe-Vonna Palmrose, Doug Prawitt, Raj Srivastava, and Hal Zeidman, and especially helpful comments by Harry Evans (editor) and two anonymous reviewers.

Editor's note: Accepted by John Harry Evans III.

Submitted: May 2010
Accepted: September 2012
Published Online: September 2012

I. INTRODUCTION

International Standard on Auditing (ISA) 600 applies to audits of *group financial statements*, where group financial statements are defined as those that include the financial information of more than one component (International Auditing and Assurance Standards Board [IAASB] 2007). This broad scope encompasses audits of financial statements of a local store owner who assembles information from two segments as well as audits of consolidated or combined financial statements of a large and complex multinational corporation. To prepare group financial statements, group management must aggregate information about components that often operate in different industries, cultures, and jurisdictions with different statutory audit requirements, accounting frameworks, and stock exchange regulations, and that require separate audits by different audit teams or different audit firms.

How much auditing is enough for audits of group financial statements? In addressing the question, ISA 600 requires the group auditor to set individual *component materiality* amounts (paragraph 21(c)).[1] These amounts allow separate component audits and determine audit work effort for each component, which ultimately determines group financial statement reliability and value. Amounts set too high expose investors to unacceptable risk that aggregate undetected component misstatement materially misstates the group financial statements (paragraph 6), while amounts set too low entail excessive audit cost. Thus, component materiality has important consequences for auditors, groups, investors, and those charged with governance, as well as other users and beneficiaries of audited group financial statements.

Despite the importance of component materiality, none of the three major auditing standards prescribes how component materiality amounts are to be determined and neither does the scholarly auditing literature. As to standards, ISA 600 merely indicates endpoints of the range of possible component materiality amounts. It says that component materiality should be "lower than the materiality level for the group financial statements as a whole" (paragraph 21(c)) and "need not be [as small as] an arithmetical portion" of group materiality (paragraph A43). As an example, for a group comprised of five equally sized components and group materiality of $100, the group auditor could set component materiality anywhere between $100 and $20 (i.e., $100/5). In the absence of more precise guidance, group audit practices vary widely and include the upper and lower limits.

Prior academic research regarding component materiality is also limited. Boritz et al. (1993) use Poisson distribution theory to aggregate audit-sampling assurance achieved across multiple components. Dutta and Graham (1998) develop a normal distribution-based method to disaggregate financial statement-level materiality to individual account balances based on investor materiality criteria for account combinations and ratios, and Turner (1997) demonstrates, via simulations, that individually immaterial misstatements can aggregate *ex post* to materially misstate key ratios. These papers do not address the group auditor's planning problem, which is to set appropriate component materiality amounts.

As to planning materiality, Messier et al. (2005, 183) note, "No research that we are aware of has investigated how planning materiality (or its allocation) . . . is handled on multi-location audits. Given the diverse nature of, and/or multinational operations of, enterprises today, research in this area is needed."[2] Glover et al. (2008a) also note the relative lack of practical guidance for

---

[1] The requirement has been adopted for U.S. private company audits (AICPA 2011) and PCAOB Auditing Standard No. 11 (PCAOB 2010) requires essentially similar "tolerable misstatement" amounts for multi-location U.S. public company audits, which parallel group audits for ISA 600. For brevity, we refer to ISA 600 alone.

[2] Similarly, Akresh et al. (1988) list materiality allocation for planning as an audit research opportunity and Blokdijk et al. (1995, 108) state that an "unresolved problem is the allocation of planning materiality and tolerable error to the specific assets, liabilities, and flows of transactions to be audited."

determining component materiality and a consequence: "Internal and peer reviews and regulatory inspections have revealed a variety of approaches [to determining component materiality, and] in some instances, reviews have discovered potentially troubling practices."[3]

As to the variety of component materiality approaches and practices, a 2006 international ISA 600 working group asked its members to describe the then current practices for determining component materiality in their respective firms and other firms. Private conversations with working group members indicated the use of ISA 600 endpoints and other practices such as one-half of group materiality for all components and setting component materiality equal to group materiality times the square root of relative component size.

In this paper, we use auditing standards, the audit risk model, prior audit research, and Bayesian probability theory to develop a *general unified assurance and materiality* model (GUAM). GUAM provides a conceptual basis for determining component materiality amounts that meet the group audit assurance objective articulated by ISA 600 while minimizing component audit cost, as well as a vehicle to explore the role of structure and group-level controls.

We begin with the group auditor's overall objective of "*reasonable assurance* about whether the [group] financial statements as a whole are *free from material misstatement*" (IAASB 2008a; emphasis added) and the (single component) *audit risk model* (ARM), which describes *audit risk* as the product of prior-to-audit *risk of material misstatement* and *detection risk* that the audit will fail to detect such misstatement when it exists. We generalize and extend the ARM within a Bayesian framework to accommodate (1) the full probability distribution of possible misstatements, (2) multiple components, and (3) group-level controls, a new element in ISA 600 and central to well-run groups and economical group audits.[4]

In the GUAM model, audit assurance is interpreted as a subjective probability distribution (Boritz et al. 1993, 230; O'Hagan and Forster 2004, 1.16) that reflects the auditor's professional beliefs about possible undetected misstatement and is denoted an *assurance profile*. We draw upon and generalize prior Bayesian audit analyses (e.g., Felix 1974; Leslie et al. 1979; Leslie 1985; Steele 1992; van Batenburg et al. 1994; Stewart 2008) that use probability distributions to express audit assurance as it relates to single entities or components.[5]

The GUAM model enables the group auditor to:

1. Represent the group auditor's overall assurance objective (e.g., achieving 95 percent assurance that total group misstatement does not exceed an exogenously predetermined group materiality of $100,000) as a target posterior (post-audit) assurance profile;
2. Disaggregate the target group posterior into target posterior assurance profiles for each component;
3. Represent the group auditor's prior (pre-audit) assurance, if any, about each component based on group-level controls and other factors pertaining to the component;
4. Derive from the target component posteriors and priors, appropriate component materiality amounts for component auditors;
5. Update the group auditor's component priors to actual posterior assurance profiles based on component auditors' findings; and finally,

---

[3] In addition, Glover et al. (2008a) suggest a computational method for determining component materiality amounts.

[4] *Group-wide controls* are defined in ISA 600 (paragraph 9(l)) as "Controls designed, implemented and maintained by group management over group financial reporting." We use the term "group-level controls" to describe controls applied by group management to one or more components, but not necessarily to all.

[5] Prior single-component Bayesian models also explored, and we incorporate, subjective inherent and control risk assessments (Kinney 1975; Bailey 1981; Jones 1999) and the results of analytical procedures (Kinney 1979; Stringer and Stewart 1996), monetary unit audit sampling, and non-statistical tests of details.

6. Aggregate the group auditor's component posteriors to form a posterior assurance profile for the group as a whole as a basis for an opinion on the group financial statements.

To assess possible implications for practice, we apply GUAM and other methods under various stylized conditions. We find that GUAM's component materiality amounts and achieved group assurance differ substantially from those of various methods used in practice. For example, for five equally sized components, group materiality of $100, and target group assurance of 95 percent, GUAM yields component materiality of $33 for each component and achieves the target 95 percent assurance, while component materiality for other methods ranges from $20 (for proportionate materiality) to $100 (for full group materiality) and achieves assurance ranges from more than 99 percent to only 18 percent, respectively.

Regarding pre-audit assurance, we find that group-level controls can facilitate substantially increased component materiality amounts, thereby reducing component auditor work. We show how GUAM can reduce aggregate component audit cost by optimizing component materiality to incorporate contextual factors, such as group structure, differing component sizes and audit costs, and component-level constraints imposed by statutory audit requirements.

We make three broad contributions. First, by generalizing and extending the single component ARM, we derive a conceptually sound, standards-based planning approach to determining component materiality. Component materiality amounts are calibrated to (1) limit, to the level acceptable by the group auditor, the audit risk that undetected component misstatements aggregate to materially misstate the group financial statements, and (2) minimize total component audit cost consistent with that overall risk level. The approach (GUAM) is amenable to practical application and provides a basis for integrating and aggregating audit results and for research on group audit methods.[6] Second, through GUAM, we show the potential impact of subgroups and group-level controls applied by group management on component audit costs and provide a framework for analytical, archival, and behavioral research on subgroups and group-level controls. Finally, our stylized application of GUAM and other methods used in practice reveals large differences in achieved assurance and component audit cost and suggests that other methods may expose group investors to excessive information risk or excessive audit cost.

Section II develops the GUAM model by generalizing the audit risk model and extending it to multiple components. Section III applies GUAM to groups with simple structures and, for stylized conditions, compares its results with those of current practice methods. Section IV extends GUAM to accommodate richer and more practical contexts, real-world constraints, and group and subgroup structures. Section V concludes and describes limitations.

II. GENERALIZING THE AUDIT RISK MODEL FOR MULTIPLE COMPONENTS

The essential innovation in GUAM is the integration of component assurance and aggregate group assurance within one Bayesian model as a basis for audit planning and forming conclusions for group financial statements. We develop the GUAM model by generalizing and extending the standard *audit risk model* (AICPA 2006; Schilder 1995; Messier et al. 2010). We initially assume a single-component "group" with a positive random variable, X, representing total undetected misstatement and an exogenously determined monetary amount, T, representing materiality for the financial statements as a whole as defined in ISA 320 (IAASB 2008b, 10). We then extend the analysis to the aggregation of multiple component assurance profiles.

[6] To encourage exploration, research, and development, Dr. Stewart (trsny@verizon.net) will provide, upon request, Excel-based software that facilitates application of the GUAM model described in this paper.

Elaborating the Audit Risk Model

The ARM is commonly expressed as:

$$AR = RMM \times DR, \tag{1}$$

where *audit risk* (*AR*), the posterior (post-audit) risk of undetected material misstatement in a single financial statement item or component, is the product of the prior (pre-audit) *risk of material misstatement* (*RMM*) and *detection risk* (*DR*). If $RMM = 50$ percent and $DR = 10$ percent, then *AR* is 5 percent. Equivalently, achieved *assurance* is $\Pr(X \leq T) = 95$ percent. *Risk of material misstatement* is defined in ISA 200, paragraph 13(n) as the "risk that the financial statements are materially misstated prior to audit." Auditors are required to assess *RMM* and use it as a basis for planning further audit procedures. *Detection risk* is the risk that *further audit procedures* (analytical procedures and substantive tests of details) will not detect the presence of material misstatement should it exist in the financial statements (IAASB 2008a).

The ARM is widely used in practice as an audit-planning tool for a single audit area or entity. Given *T*, a target *AR* level, and assessed *RMM*, further audit procedures should be planned that limit detection risk to:

$$DR = AR/RMM. \tag{2}$$

For example, if the target $AR = 5$ percent and $RMM = 50$ percent, then $DR = 5$ percent/50 percent = 10 percent.[7]

In the Bayesian GUAM model, the ARM's prior and posterior *probabilities*, *RMM* and *AR*, are replaced by prior and posterior probability *distributions*, respectively. The probability *DR* is replaced by the *likelihood function* induced by evidence from further audit procedures. Under GUAM, prior assurance is updated for new evidence to determine posterior assurance according to *Bayes' rule* (see Appendix A), which generalizes the simple multiplicative Equation (1).

About GUAM and Gamma Distributions

We represent priors, posteriors, and likelihoods with gamma probability distributions for two broad reasons. First, gamma distributions are already implicit in auditing practice via the ARM. The ARM yields proper conclusions *if and only if* its probabilities *RMM*, *DR*, and *AR* are from exponential distributions, which are well-known members of the gamma family (Stewart 2012). Exponential assurance profiles arise when there is no indication of misstatement. Their mode is zero and they decline exponentially as *X* increases, indicating an increasingly remote probability of large misstatement. When an exponential distribution is used to represent prior assurance, *RMM* is the probability represented by the area under the probability distribution for *X* greater than materiality, *T*. Further, gamma likelihood functions are embedded in practice through monetary unit sampling (MUS) and it is often possible to represent the results of other audit procedures as evidentially equivalent MUS samples (Smith 1976; Leslie et al. 1979; Steele 1992).

The second reason for using gamma distributions is that they have three valuable statistical properties. Given a gamma prior and a gamma likelihood, Bayes' rule results in a posterior distribution that is also a gamma (i.e., the gamma is a *conjugate* prior), which keeps the model within the gamma family. Also, if there is no indication of misstatement, then the exponential distribution provides the most conservative representation of prior assurance (and *RMM*) in the

[7] The ARM has been criticized as not adequately representing audit assurance as an essentially Bayesian process (Leslie 1984; Kinney 1984, 1989; Akresh 2010) and it has no construct for the combined risk of misstatement of multiple components (Kinney 1993). Also, as planning models, neither the ARM nor GUAM incorporates the risk of incorrect rejection of materially correct information that arises when audits do not go as planned.

sense that any competing probability distribution requires additional restrictive assumptions. The exponential distribution's status as a *maximum entropy* distribution (O'Hagan and Forster 2004, 5.42) ensures that it is maximally noncommittal with regard to missing information (Jaynes 1957). Finally, gamma-distributed component misstatements aggregate to group misstatement and the distribution of the aggregate can be approximated by another gamma distribution, which again keeps the entire component/group model in the gamma family.

Gamma distributions have two parameters, a *shape* parameter $\alpha > 0$, which determines the basic shape of the distribution and a *scale* parameter $\beta > 0$, which for any given shape determines the *diffusion* or *spread* of X. A gamma distribution with $\alpha = 1$ is the familiar *exponential* distribution. We use the following notation for a gamma-distributed random variable X:

- $g(x;\alpha,\beta)$ denotes the *probability density* function.
- $G(x;\alpha,\beta)$ denotes the *cumulative distribution* function. That is, $G(x;\alpha,\beta) = \Pr(X \leq x)$, which equals the relative area under the $g(x;\alpha,\beta)$ curve between 0 and x.
- $Q(P;\alpha,\beta)$ denotes the *percentile function* and is the inverse of the cumulative distribution function. If $G(x;\alpha,\beta) = P$, then $x = Q(P;\alpha,\beta)$ is the 100Pth *percentile* of X, which is convenient to write as $\beta Q_P(\alpha)$.

Appendix A provides a brief summary of gamma distribution properties and their Bayesian interpretation, including a graphical depiction of various gamma distributions.

When prior assurance or *RMM* is assessed in practice there is usually no indication that misstatement exists, and it is commonly assumed that the modal misstatement is zero and larger amounts of misstatement are increasingly improbable. These characteristics describe the essence of the exponential distribution. Another common assumption is that planned further audit procedures will not indicate misstatement and will also induce mode zero exponential likelihood functions. As mentioned above, these assumptions are implicit in the ARM in that *RMM* and *DR* are from exponential distributions. We assume, therefore, that component priors and anticipated likelihood functions can be represented by exponential ($\alpha = 1$) distributions.

In the case of exponential distributions, if P is the probability that total misstatement X does not exceed some amount x (i.e., $\Pr(X \leq x) = G(x;1,\beta) = P$), then the 100$P$th percentile is:

$$\beta Q_P(1) = -\beta \ln(1 - P). \tag{3}$$

In audit applications, the factor $Q_P(1)$ is sometimes called a *reliability factor* (or a *confidence factor* [AICPA 2008]) and we designate it R_P, thus:

$$R_P = Q_P(1) = -\ln(1 - P). \tag{4}$$

For example, $R_{0.95} = -\ln(1 - 0.95) = -\ln(0.05) = 3.0$. Other commonly encountered R_p factors include $R_{0.90} = 2.3$, $R_{0.86} = 2.0$, $R_{0.70} = 1.2$, $R_{0.63} = 1.0$, and $R_{0.50} = 0.7$.

Equations (3) and (4) are particularly useful for P assessed at T, thus representing the probability that X does not exceed materiality T, because they relate the ARM to a Bayesian interpretation. The probability P that misstatement does not exceed T can also be represented by an assurance profile with 100Pth percentile equal to T. If there is no indication of misstatement, then the appropriate assurance profile is the exponential $g(x;1,\beta)$, where from Equations (3) and (4):

$$\beta = \frac{T}{R_P}. \tag{5}$$

Therefore, if risk of material misstatement is assessed to be *RMM*, then the appropriate Bayesian prior distribution is $g(x;1,\beta)$, where:

$$\beta = \frac{T}{R_{1-RMM}}. \tag{6}$$

For example, if $RMM = 50$ percent, then $R_{1-RMM} = R_{0.50} = 0.7$ and the prior is $g(x;1,T/0.7)$. The probabilities DR and AR in the ARM can be similarly translated into unique exponential assurance profiles via Equations (4) and (5). If desired $AR = 5$ percent, then $R_{1-AR} = R_{0.95} = 3.0$ and the target posterior assurance profile is $g(x;1,T/3.0)$. Finally, based on these examples, target detection risk is DR = 5 percent/50percent = 10 percent, and $R_{1-DR} = R_{0.90} = 2.3$, which is represented with the likelihood function $g(x;1,T/2.3)$. Thus, the ARM model, $AR = RMM \times DR$, is replaced by the Bayesian model:

$$\underbrace{g\left(x;1,\frac{T}{R_{1-AR}}\right)}_{\text{Posterior Density}} = \underbrace{g\left(x;1,\frac{T}{R_{1-RMM}}\right)}_{\text{Prior Density}} \times \underbrace{g\left(x;1,\frac{T}{R_{1-DR}}\right)}_{\text{Likelihood Function}} \times \underbrace{(\beta + B)}_{\text{Constant}}. \tag{7}$$

That is, the *posterior* equals the *prior* times the *likelihood*, times a *constant* to ensure total posterior probability equals 1 (see Appendix A). For example, the ARM formulation 5 percent = 50 percent × 10 percent is replaced by:

$$g\left(x;1,\frac{T}{3.0}\right) = g\left(x;1,\frac{T}{0.7}\right) \times g\left(x;1,\frac{T}{2.3}\right) \times \left(\frac{T}{0.7} + \frac{T}{2.3}\right).$$

The expression T/R_{1-DR} appears in the monetary unit sampling (MUS) literature as a determinant of sample size (AICPA 2008, 6.24, C.2), where DR is sampling risk. If Y is population size, then the appropriate MUS sample size is:[8]

$$n = \frac{Y}{T/R_{1-DR}}. \tag{8}$$

If no misstatements are found, then the induced likelihood function has a mode of zero and is essentially the exponential distribution $g(x;1,T/R_{1-DR})$.

Using Bayes' Rule for Audit Planning

When an exponential prior assurance profile $g(x;1,\beta)$ is updated by an exponential likelihood function $g(x;1,B)$ according to Bayes' rule, the resulting posterior assurance profile is the exponential assurance profile $g(x;1,\beta')$, where:

$$\beta' = \frac{1}{1/\beta + 1/B}. \tag{9}$$

(see Appendix A). Building on the previous example, if the auditor's prior is $g(x;1,T/0.7)$ (equivalent to $RMM = 50$ percent) and the likelihood function is $g(x;1,T/2.3)$ (equivalent to $DR = 10$ percent), then from Equation (9):

$$\beta' = 1/\left(\frac{1}{T/0.7} + \frac{1}{T/2.3}\right) = \frac{T}{3.0},$$

and the posterior assurance profile is $g(x;1,T/3.0)$ (equivalent to $AR = 5$ percent).

[8] For example, if $DR = 10$ percent, $T = 100$, and $Y = 10{,}000$, then $R_{1-DR} = R_{0.90} = 2.3$, and $n = 10{,}000/(100/2.3) = 230$.

For audit planning, Bayes' rule is applied in reverse, analogous to the ARM in Equation (2). If the auditor desires a posterior assurance profile $g(x;1,\beta')$ and starts with a prior $g(x;1,\beta)$, where $\beta' < \beta$, then from Equation (9) the required likelihood function is $g(x;1,B)$, where:

$$B = \frac{1}{1/\beta' - 1/\beta}. \tag{10}$$

For example, if the desired posterior is $g(x;1,T/3.0)$ and the prior is $g(x;1,T/0.7)$, then from Equation (10) the target likelihood function from further audit procedures is $g(x;1,T/2.3)$.[9]

When auditors assess pre-audit risk of material misstatement as a certainty (RMM = 100 percent), their entire assurance must come from further audit procedures and audit risk equals detection risk (AR = 100 percent $\times$ DR = DR). In Bayesian terms, having no basis for prior assurance (denoted *negligible* prior assurance) is represented as a "flat" exponential distribution $g(x;1,\beta)$ with a scale parameter β so large relative to T that $1/\beta$ can be treated as zero in Equations (9) and (10), and consequently, $g(x;1,\beta') = g(x;1,B)$. The net effect of a negligible prior is that essentially all the auditor's assurance comes from further audit procedures and, thus, a negligible prior achieves the same result as setting RMM = 100 percent.

Extending the Model to Groups

Having obtained a component assurance profile $g(x_i;\alpha_i,\beta_i)$ for each component i, $i = 1,\ldots,N$, the group auditor can determine the group assurance profile—the probability distribution of the aggregate random variable $\sum_{i=1}^{N} X_i$ across all components. For example, if there are ten components, then the auditor is interested in the potential misstatement aggregated across all ten components in the group, or the probability distribution of $X_1 + \cdots + X_{10}$. To accommodate multiple components, we add the following assumptions and notation:

- Component materiality for Component i is T_i, which is used by the auditor of Component i to plan and perform the audit of the component.
- There is a consistent measure of relative component size across the group, denoted Y_i (e.g., total revenues or total assets).
- The component potential misstatement amounts, the X_i's, are stochastically independent.[10]

Given that the X_i's are independent, the distribution of their aggregate is formed by the *convolution* of the N component probability distributions (DeGroot and Schervish 2002, 172), which we denote the *aggregate* distribution. In general, the aggregate of a set of gamma distributions $g(x_i;\alpha_i,\beta_i)$ is a complicated probability distribution that can be approximated with a gamma distribution having the appropriate expected value and variance (Stewart et al. 2007). Because $E(X_i) = \alpha_i\beta_i$ and $\mathrm{Var}(X_i) = \alpha_i\beta_i^2$ and the X_i are independent, the expected value and variance of the aggregate are:

---

[9] As mentioned above, if the target likelihood function is $g(x;1,B)$ with $B = T/R_{1-DR}$ and MUS is used, then sample size is $n = Y/B = Y/(T/R_{1-DR})$. Thus, $1/B = R_{1-DR}/T$ is the required average sampling rate, or, equivalently, $B = T/R_{1-DR}$ is the required average sampling interval.

[10] Under the notion of subjective probability, two random variables are independent if learning the value of one does not change one's belief about the other (O'Hagan and Forster 2004, 4.5). For separately managed components and absent any *ex ante* belief that systemic management fraud or other misstatement is present, independence is a reasonable planning assumption and is assumed in other relevant literature such as Zuber et al. (1983), Boritz et al. (1993), Dutta and Graham (1998), and Glover et al. (2008a, 2008b).

$$E\left(\sum X_i\right) = \sum E(X_i) = \sum \alpha_i\beta_i \quad \text{and} \quad \mathrm{Var}\left(\sum X_i\right) = \sum \mathrm{Var}(X_i) = \sum \alpha_i\beta_i^2,$$

and $g(\sum x_i; \alpha, \beta)$ is the approximating gamma distribution, where:

$$\alpha = \frac{\left[E\left(\sum X_i\right)\right]^2}{\mathrm{Var}\left(\sum X_i\right)} = \frac{\left(\sum \alpha_i\beta_i\right)^2}{\sum \alpha_i\beta_i^2}, \tag{11}$$

$$\beta = \frac{\mathrm{Var}\left(\sum X_i\right)}{E\left(\sum X_i\right)} = \frac{\sum \alpha_i\beta_i^2}{\sum \alpha_i\beta_i}. \tag{12}$$

(see Appendix A). In the general case of unequal β_i's, the approximation is good enough for any group likely to be encountered in practice. For example, Stewart et al. (2007) show that if the approximation indicates that 95 percent assurance has been achieved, then actual assurance lies between 94.9 percent and 95 percent. In the special case where $\beta_1 = \cdots = \beta_N$, the gamma $g(\sum x_i; \alpha, \beta)$ with $\alpha = \sum \alpha_i$ and $\beta = \beta_i$ for $i = 1,\ldots,N$ is the correct aggregate distribution, so there is no approximation error (DeGroot and Schervish 2002, 298).

For N identical exponential component distributions $g(x_i;1,\beta)$, the aggregate distribution further reduces to $g(\sum x_i; N, \beta)$. For example, Figure 1, Panel A depicts two identical exponential component assurance profiles $g(x_i;1,100/3.0)$ whose aggregate results in the group assurance profile $g(x_1 + x_2;2,100/3.0)$ on the right.[11] The two component posterior profiles could result from each component auditor applying component materiality of $T_i = 100$ (i.e., $T_1 = T_2 = T$, the upper limit indicated by ISA 600). The 95th percentile marked on the x axis is $\beta Q_{0.95}(1) = (100/3.0) \times 3.0 = 100$ for each component and the 95th percentile is $\beta Q_{0.95}(2) = (100/3.0) \times 4.74 = 158$ for the group. Because the group 95th percentile (158) exceeds $T = 100$, the component assurance profiles in Panel A do not meet the group auditor's assurance objective. Specifically, there is only 80 percent assurance that total misstatement does not exceed $T = 100$ (i.e., $\Pr(X_1 + X_2 \leq 100) = G(100;2,100/3.0) = 80$ percent).

In contrast, Figure 1, Panel B shows that the group objective *is* achieved if sufficient auditing is performed to reduce the component scale parameters from $\beta = 100/3.0$ to $\beta = 100/4.74$ (e.g., by increasing the MUS sample size from 300 to 474). This can be done by instructing the component auditors to use $T_1 = T_2 = (100/4.74) \times 3.0 = 63$, which are equal to the respective 95th percentiles. If the component auditors achieve $\Pr(X_1 \leq 63) = 95$ percent and $\Pr(X_2 \leq 63) = 95$ percent, respectively, then the group auditor achieves the $\Pr(X_1 + X_2 \leq 100) = 95$ percent group audit objective.

III. GUAM FOR GROUPS WITH SIMPLE STRUCTURES

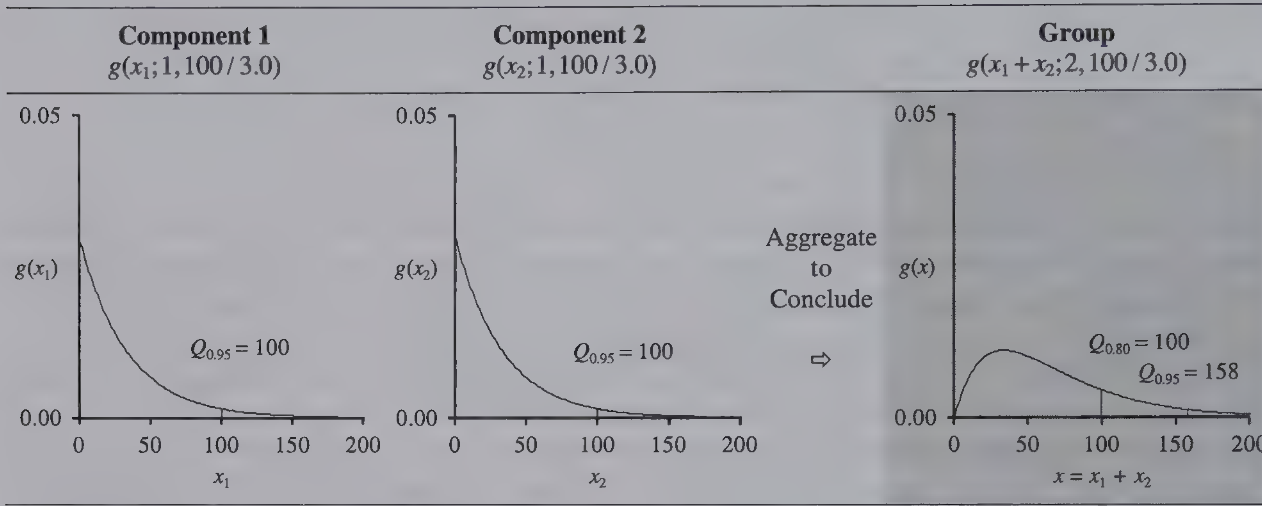
In this section we use the general unified assurance and materiality (GUAM) model to derive a method for calculating component materiality amounts for groups consisting of separately operated and audited stand-alone units. *Achieved group assurance* (AGA) is the group auditor's assurance that the aggregate of undetected misstatements in the group financial statements does not exceed group materiality, that is, $AGA = \Pr(\sum X_i \leq T)$. For specificity, we assume that the group audit assurance objective is to achieve $AGA = 95$ percent.

---

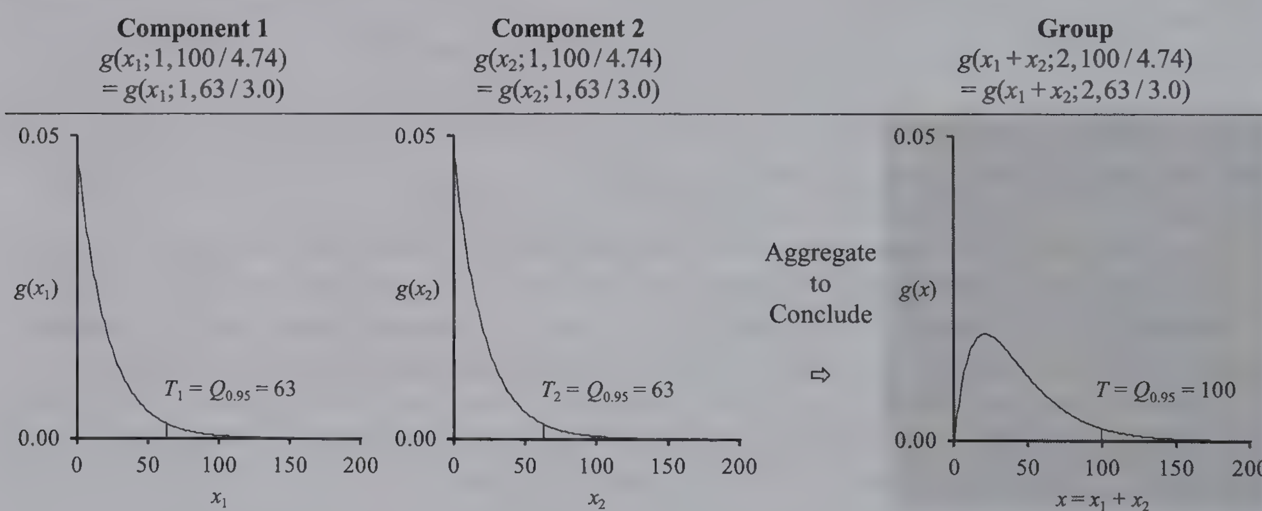
[11] It may appear counter-intuitive that component assurance profiles with a modal misstatement of zero should aggregate to a group profile with mode > 0. The analogy of a portfolio of insurance policies may be helpful. The modal claim for each individual policy in a particular year is likely to be zero, but the modal claim for a portfolio comprised of many such policies would be greater than zero.

FIGURE 1
Aggregating Component Assurance Profiles to Form the Group Assurance Profile

Panel A: Two Exponential Distributions with $\beta = T/3.0$; Group Assurance Achieved Relative to $T = 100$ is 80 Percent



Panel B: Two Exponential Distributions with $\beta = T/4.74$; Group Assurance Achieved Relative to $T = 100$ is 95 Percent



The aggregation of component assurance profiles is represented by the aggregate of their probability distributions. The random variables X_1 and X_2 are each distributed $g(x_i; 1, \beta)$. Their sum is distributed $g(x_1 + x_2; 2, \beta)$, which is the group assurance profile. The component distributions in Panel A have $\beta = 100/3.0$. Their aggregate has 95th percentile $\beta Q_{0.95}(2) = 4.74\beta = 158$, which exceeds $T = 100$ and thus does not achieve the group assurance objective. Reducing the component scale parameters to $\beta = 100/4.74 = 63/3.0$ as in Panel B achieves the group audit objective. Component materiality in Panel B is set equal to the 95th percentile of the component assurance profile giving $T_1, T_2 = \beta R_{0.95} = (63/3.0) \times 3.0 = 63$.

GUAM Calculation of Component Materiality

To achieve $AGA = 95$ percent, the group auditor must obtain component posterior assurance profiles that aggregate to a posterior group assurance profile $g(\sum x_i; \alpha', \beta')$ that has a 95th percentile equal to group materiality T. That is:

$$\beta' Q_{0.95}(\alpha') = T, \tag{13}$$

where the group parameters α' and β' are derived from the component posteriors via Equations (11) and (12), respectively.

The key to GUAM is determining component posteriors that aggregate to a group posterior that satisfies Equation (13). Equation (13) can be satisfied by assigning a positive *weight*, w_i, to each component, $i = 1,\ldots,N$, such that the weights sum to 1, and setting:

$$\beta_i' = \frac{T}{Q_{0.95}(1/w_i)} \tag{14}$$

for each target component posterior $g(x_i;1,\beta_i')$. If the components are equally weighted, then Equation (13) is met exactly and $AGA = 95$ percent. If the components are not equally weighted, then Equation (13) is met approximately and $AGA \approx 95$ percent.[12] Because *any* combination of positive weights summing to 1 will result in a set of component posteriors that achieves $AGA \approx 95$ percent, weights can be optimized to achieve secondary objectives such as cost minimization as discussed later in this section and again in Section IV.

If the group auditor's prior component assurance profile for component i is $g(x_i;1,\beta_i)$, and $\beta_i' < \beta_i$, then from Equation (10) the target likelihood function required to update the prior into the target posterior is $g(x_i;1,B_i)$, where:

$$B_i = \frac{1}{1/\beta_i' - 1/\beta_i}. \tag{15}$$

By construction, achieving target likelihood functions $g(x_i;1,B_i)$ for components $i = 1,\ldots,N$ will achieve the group assurance objective.

The *group* auditor's Bayesian perspective is that the *component* auditor must deliver the target likelihood function, $g(x_i;1,B_i)$. The *component* auditor's Bayesian perspective is that $g(x_i;1,B_i)$ is the target posterior assurance profile for the component audit. The component objective can be communicated by the group auditor instructing the component auditor to use component materiality equal to the 95th percentile of $g(x_i;1,B_i)$, namely:

$$T_i = B_i Q_{0.95}(1) = B_i R_{0.95}. \tag{16}$$

In ARM terms, use of Equation (16) is analogous to instructing the component auditor to use T_i to achieve $AR = 5$ percent, as illustrated in the earlier example related to Equation (1). Computationally, the formula for component materiality follows from Equations (14), (15), and (16):

$$T_i = \frac{TR_{0.95}}{Q_{0.95}(1/w_i) - T/\beta_i}, \tag{17}$$

where $T/\beta_i < Q_{0.95}(1/w_i)$.

---

[12] Extensive simulations show that the approximation error in Equation (13) is small enough that it can be ignored for any group likely to be encountered in practice (Stewart 2012).

If the group auditor has negligible prior component assurance (or $RMM_i = 100$ percent), then $1/\beta_i \approx 0$ and Equation (17) simplifies to:

$$T_i = \frac{TR_{0.95}}{Q_{0.95}(1/w_i)}. \tag{18}$$

If the N components are equally weighted, then $w_i = 1/N$ and Equation (18) reduces to:

$$T_i = \frac{TR_{0.95}}{Q_{0.95}(N)}. \tag{19}$$

Therefore, in this simplest form, T_i is T times the ratio of the 95th percentile of the target component posterior ($\beta' R_{0.95}$) to the 95th percentile of the target group posterior ($\beta' Q_{0.95}(N)$).[13]

If the group auditor can independently establish a non-negligible prior on Component i (i.e., $RMM_i < 100$ percent) based on assessing inherent risk of material misstatement for the component, evaluating and testing group-level controls that apply to the component, and/or applying analytical procedures at the group level, then from Equations (4) and (6) the component prior is $g(x;1,\beta_i)$ where $\beta_i = T_i/R_{1-RMM_i}$. Making this substitution for β_i in Equation (17) and solving for T_i gives:

$$T_i = \frac{TR_{0.95}}{Q_{0.95}(1/w_i)}\left(1 + \frac{R_{1-RMM_i}}{R_{0.95}}\right), \tag{20}$$

or T_i equals T_i for a negligible prior (Equation (18)) times a factor $1 + (R_{1-RMM_i}/R_{0.95})$that reflects the group auditor's prior assurance for Component i. As the group auditor's prior assurance for Component i increases, so does R_{1-RMM_i} and T_i increases accordingly, reflecting that less work and less assurance is required from the component auditor.

Figure 2 illustrates the determination of T_i values for a group with two equally weighted components, $(w_1,w_2) = (0.5,0.5)$, under different component prior scenarios. In Panel A the group auditor has negligible prior assurance for both components. From Equation (19):

$$T_i = \frac{TR_{0.95}}{Q_{0.95}(2)} = \frac{100 \times 3.0}{4.74} = 63,$$

and $(T_1,T_2) = (63,63)$. In Panel B the group auditor rates RMM_1 as 50 percent for Component 1, so that from (4), $R_{0.50} = 0.7$, and T_1 can be increased in accordance with Equation (20) to:

$$T_1 = \frac{TR_{0.95}}{Q_{0.95}(2)}\left(1 + \frac{R_{1-RMM_1}}{R_{0.95}}\right) = \frac{100 \times 3.0}{4.74}\left(1 + \frac{0.7}{3.0}\right) = 63 \times 1.23 = 78.$$

Therefore, $(T_1,T_2) = (78,63)$. In both scenarios, the group auditor's posterior assurance profiles for the components are $g(x_i;1,63/3.0)$, $i = 1,2$. Similarly, the group posteriors are $g(x_1 + x_2;\alpha',\beta') = g(x_1 + x_2;2,63/3.0)$ for both scenarios and the group 95th percentiles are:

$$\beta' Q_{0.95}(\alpha') = \left(\frac{63}{3.0}\right) \times 4.74 = 100,$$

which equals T as required.

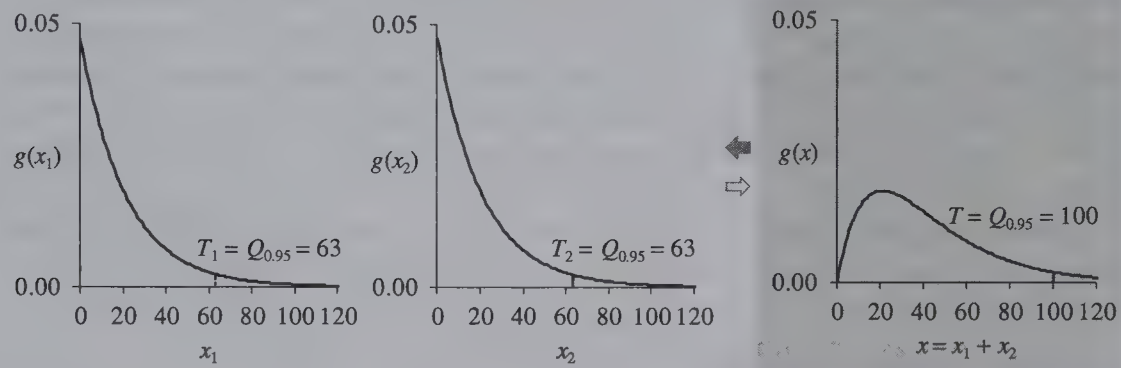
Thus, while the routes to forming group assurance in Figure 2 differ, the resulting posteriors are identical, reflecting the fact that a group auditor with a knowledge base that allows a non-negligible

[13] For example, the $N = 2$ group in Figure 1, Panel B, shows $T_i = T \times R_{0.95}/Q_{0.95}(2) = 100 \times 3.0/4.74 = 63$.

FIGURE 2
Determining Component Materiality with Negligible and Moderate Group Auditor Prior Assurance for Component Misstatement

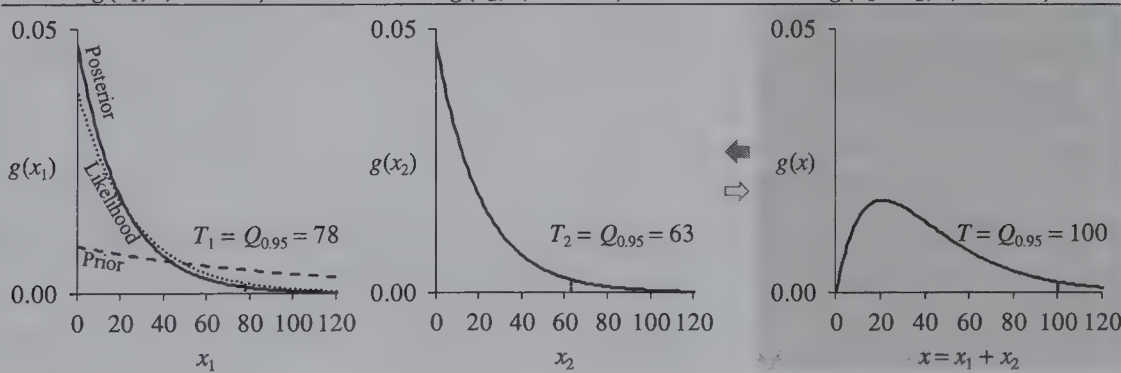
Panel A: Negligible Group Auditor Priors for Component Misstatement

	Component 1	Component 2	Group
Prior	Negligible	Negligible	
Likelihood	$g(x_1;1,63/3.0)$	$g(x_2;1,63/3.0)$	
Posterior	$g(x_1;1,100/4.74) =$ $g(x_1;1,63/3.0)$	$g(x_2;1,100/4.74) =$ $g(x_2;1,63/3.0)$	$g(x_1+x_2;2,100/4.74) =$ $g(x_1+x_2;2,63/3.0)$



Panel B: Moderate Group Auditor Prior for Component 1 Misstatement Based on Assessment of Applicable Group-Level Controls

	Component 1	Component 2	Group
Prior	$g(x_1;1,78/0.7)$	Negligible	
Likelihood	$g(x_1;1,78/3.0)$	$g(x_2;1,63/3.0)$	
Posterior	$g(x_1;1,78/3.7) =$ $g(x_1;1,63/3.0)$	$g(x_2;1,100/4.74) =$ $g(x_2;1,63/3.0)$	$g(x_1+x_2;2,100/4.74) =$ $g(x_1+x_2;2,63/3.0)$



← *Planning* works from right to left: (1) target component posteriors are determined based on the group assurance objective (95th percentile equal to $T = 100$), and (2) target component likelihood functions are determined from those posteriors and the component priors.

→ *Concluding* works from left to right: (1) the group auditor's component posteriors are derived from component priors and likelihoods, and (2) the component posteriors are aggregated to derive the group posterior.

In Panel A, component priors are negligible and therefore the target likelihoods and posteriors coincide. The component materialities are $(T_1,T_2) = (63,63)$. In Panel B, the group auditor's prior for Component 1 is $g(x_1;1,78/0.7)$ and the target likelihood function is $g(x_1;1,78/3.0)$, which has a 95th percentile of 78. Accordingly, the component materialities are $(T_1,T_2) = (78,63)$.

component prior can apply Equation (20) to increase T_i and reduce the required component auditor effort. While little is known about current practices in evaluating group-level controls, optimal group audit design requires trading off the cost of developing the group auditor's component priors against the potential savings in component auditor effort.

Group Assurance and Cost Metrics

As indicated in Section I, the component materiality amounts $(T_1,\ldots,T_N)$ that the group auditor sets directly determine how much auditing is performed and, thus, affect the quality and cost of group audits. In this section we develop metrics for the achieved group assurance (*AGA*) and relative cost that result from any component materiality amounts.

Regarding group audit quality as measured by *AGA*, in the case of GUAM, $(T_1,\ldots,T_N)$ is constructed to ensure $AGA \approx 95$ percent, but application of alternative component materiality methods may yield greater or lesser *AGA* for the same conditions. For comparisons across methods, we assume that the group auditor's posterior component assurance profile for each component is an exponential distribution $g(x_i;1,\beta_i')$. Aggregating these component posteriors via Equations (11) and (12) gives the group posterior $g(\sum x_i;\alpha',\beta')$, from which achieved group assurance is $AGA = G(T;\alpha',\beta')$. If component priors are negligible and component audits are conducted to achieve 95 percent assurance that $X_i \leq T_i$, then $T_i = \beta_i' R_{0.95}$, therefore $\beta_i' = T_i/R_{0.95}$, and from Equations (11) and (12):

$$AGA = G(T;\alpha',\beta'), \quad \text{with} \quad \alpha' = \frac{\left(\sum T_i\right)^2}{\sum T_i^2} \quad \text{and} \quad \beta' = \frac{\sum T_i^2}{R_{0.95}\sum T_i}. \tag{21}$$

Equation (21) requires only $(T_1,\ldots,T_N)$ and that the group auditor's posterior component assurance profiles are $g(x_i;1,\beta_i')$, and thus provides a way to calculate *AGA* across methods. If $T_1 = \cdots = T_N$, then Equation (21) reduces to $AGA = G(T;N,\beta')$, where $\beta' = T_1/R_{0.95} = \cdots = T_N/R_{0.95}$. For example, if $T = 100$ is allocated equally to two components so that $(T_1,T_2) = (50,50)$, then $\alpha' = N = 2$, $\beta' = 50/3.0$, and $AGA = G(100;2,50/3.0) = 98$ percent.

To be specific regarding cost, we assume that component audits entail a single substantive test that is equivalent in information and cost to a misstatement-free MUS sample. That is, we assume that component audit variable cost is proportional to the appropriate MUS sample size. From Equation (8), the sample size for $i = 1,\ldots,N$, is:

$$n_i = \frac{Y_i}{T_i/R_{0.95}}.$$

If the *variable cost per unit* of auditing items in Component i is denoted VCU_i, then the *total variable cost* for component i (denoted TVC_i) is proportional to:

$$TVC_i = VCU_i \times n_i = VCU_i\left(\frac{Y_i}{T_i/R_{0.95}}\right) = \frac{VCU_i \times Y_i}{T_i/R_{0.95}}.$$

If the VCU_i are equal across components, then the *total variable cost* of auditing all N components simplifies to:[14]

$$TVC = \sum TVC_i = \sum \frac{Y_i}{T_i/R_{0.95}}. \tag{22}$$

By comparison, if the entire group (of size $\sum Y_i$) was treated as a "virtual single entity" and each component audited using group materiality T (i.e., setting $T_i = T$ for all i), then total variable

[14] In contexts where VCU_i varies across components, we replace Y_i with $VCU_i \times Y_i$.

cost for the group would be minimized and would equal $(\sum Y_i)/(T/R_{0.95})$. Using this notional minimum cost for a virtual single entity as the divisor, *TVC* can be converted to a *relative total variable cost* index (*RTVC*):

$$RTVC = \frac{\sum Y_i/(T_i/R_{0.95})}{\left(\sum Y_i\right)/(T/R_{0.95})} = \frac{\sum Y_i/T_i}{\left(\sum Y_i\right)/T}. \tag{23}$$

If $T_1 = \cdots = T_N$, then *RTVC* is simply the ratio of T to (constant) T_i, that is $RTVC = T/T_i$. Because *RTVC* does not depend on the method used to determine $(T_1,\ldots,T_N)$, Equation (23) provides a way to compare the *RTVC* results across different component materiality methods.

To illustrate application of Equation (23), consider the GUAM-based $Y_1 = Y_2$ scenarios in Figure 2 with respect to total variable cost given different component priors. Panel A assumed negligible group auditor priors for both components so $(T_1,T_2) = (63,63)$ and $RTVC = 100/63 = 1.58$, or 58 percent more than if T were used throughout, that is, $(T_1,T_2) = (100,100)$. Panel B assumed a moderate group auditor prior for Component 1, so $(T_1,T_2) = (78,63)$ and:

$$RTVC = \frac{1/78 + 1/63}{2/100} = 1.43,$$

or 43 percent more than if T were used throughout. Thus, incorporating even a moderate group-level prior for one component allows about a 10 percent reduction in this group's *RTVC*.

Comparison of Alternative Component Materiality Methods

We compare GUAM with alternative component materiality methods by measuring *AGA* (Equation (21)), and *RTVC* (Equation (23)). We compare groups with equally sized stand-alone components and we assume negligible component priors.[15]

The methods illustrated include the endpoints for component materiality indicated by ISA 600 and three other methods used in global network firm practice today. The lower limit in ISA 600, denoted PROP, allocates T to components in proportion to their size; the upper limit, denoted FULL, sets $T_i = T$. As to other practice methods, MACM allocates *maximum aggregate component materiality* (a tabulated multiple of T) to components in proportion to the square root of size (Glover et al. 2008a); SQRT sets T_i equal to T times the square root of the relative size of the component (Zuber et al. 1983); and HALF sets $T_i = T/2$ regardless of the number of components or component size. The formulas for T_i and *RTVC* under each of the methods are summarized in Table 1.

Like GUAM, MACM and SQRT are based on underlying probability models. For MACM, the key factor is the multiple m (see Table 1), which is determined probabilistically based on independent Bernoulli trials and the binomial distribution (Glover et al. 2008b).[16] The SQRT method is based on the X_i being independent normally distributed random variables with mean zero and variance proportional to Y_i.[17]

---

[15] GUAM explicitly accounts for components of unequal sizes, non-negligible priors, and other factors, but the other methods do not and the lack of adaptability limits our ability to provide other contextual comparisons.

[16] The components are regarded as N independent pass/fail Bernoulli trials in which each outcome is dichotomously either zero misstatement or misstatement exactly equal to some constant material amount, and m follows from the binomial distribution. As the authors note, the MACM model "does not explicitly consider the ... risk that multiple components might contain undetected misstatements that are immaterial at the component level but aggregate to an amount that is material at the group level, or the need for redistribution of materiality between components due to statutory audits or substantial variability in the size of components" (Glover et al. 2008b).

[17] $\mathrm{Var}(X_i) = \sigma^2 Y_i$ and, because the X_i are independent, $\mathrm{Var}(\sum X_i) = \sigma^2 \sum Y_i$. As with GUAM, T and T_i are 95th percentiles. For normal distributions, $T_i = u_{0.95}\sigma\sqrt{Y_i}$ where $u_{0.95}$ is the 95th percentile of the standard normal deviate. Similarly, $T = u_{0.95}\sigma\sqrt{\sum Y_i}$. Therefore, $T_i/T = (u_{0.95}\sigma\sqrt{Y_i})/(u_{0.95}\sigma\sqrt{\sum Y_j})$, and $T_i = T\sqrt{Y_i/\sum Y_j}$.

TABLE 1

Summary of T_i and $RTVC$ Formulas for Component Materiality Methods N Components, Negligible Group Auditor Priors

Panel A: Formulas for Component Materiality Methods

	Component Materiality, T_i		Relative Total Variable Cost, $RTVC$[a]	
Method	**General Y_i**	**$Y_1 = \cdots = Y_N$**	**General Y_i**	**$Y_1 = \cdots = Y_N$**
GUAM	$\frac{TR_{0.95}}{Q_{0.95}(1/w_i)}, \sum w_i = 1$	$\frac{TR_{0.95}}{Q_{0.95}(N)}$	$\frac{\sum Y_i Q_{0.95}(1/w_i)}{\sum Y_i R_{0.95}}$	$\frac{Q_{0.95}(N)}{R_{0.95}}$
MACM[b]	$mT\frac{\sqrt{Y_i}}{\sum\sqrt{Y_j}}$	$\frac{mT}{N}$	$\frac{\left(\sum\sqrt{Y_i}\right)^2}{m\sum Y_i}$	$\frac{N}{m}$
SQRT	$T\sqrt{\frac{Y_i}{\sum Y_j}}$	$\frac{T}{\sqrt{N}}$	$\frac{\sum\sqrt{Y_i}}{\sqrt{\sum Y_i}}$	$\sqrt{N}$
PROP	$T\frac{Y_i}{\sum Y_j}$	$\frac{T}{N}$	N	N
HALF	$\frac{T}{2}$	$\frac{T}{2}$	2	2
FULL	T	T	1	1

Negligible priors are assumed. General formulas for unconstrained component sizes and formulas for equally sized components are both given.

[a] Per Equation (23).

[b] From Glover et al. (2008a), a tabulated multiple m is applied to T giving *maximum aggregate component materiality* (MACM), which is allocated in proportion to the square root of size. See Panel B.

Panel B: MACM Multiples

N	2	3–4	5–6	7–9	10–14	15–19	20–25	26–30
m	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
N	31–40	41–50	51–64	65–80	81–94	95–110	111–130	131+
m	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0

Table 2 compares T_i, $RTVC$, and AGA for two to ten equally sized components for the various component materiality methods assuming negligible prior assurance and Figure 3 depicts results graphically. For the assumed conditions, GUAM is the only method for which $AGA = 95$ percent is consistently achieved. PROP consistently achieves AGA $>$ 95 percent and HALF achieves it only for $N = 2$, while the other alternatives are always ineffective because $AGA <$ 95 percent. GUAM has higher $RVTC$ than the ineffective alternatives and has lower $RVTC$ than the effective PROP (and HALF for $N = 2$). Overall, GUAM is consistently effective and achieves target effectiveness at lower relative cost than other effective methods under these conditions.[18]

[18] In extended untabulated results, the patterns evident for two to ten components continue to hold. For example, with respect to AGA, the plot for GUAM remains constant at 95 percent, while plots for MACM, SQRT, FULL, and HALF continue to decline.

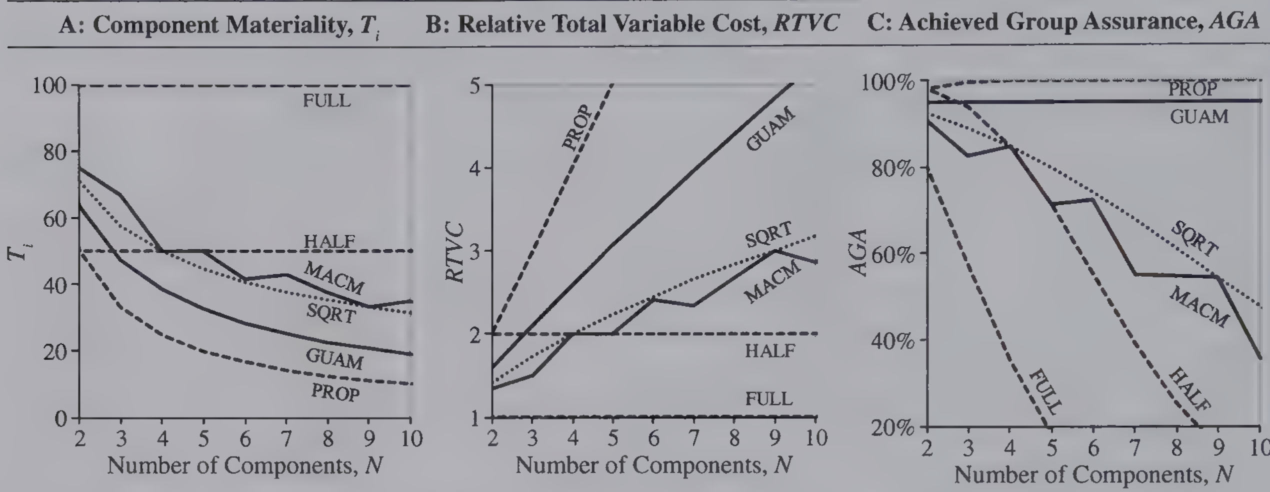
TABLE 2

T_i, $RTVC$, and AGA for $N = 2$ to 10 Equally Weighted Components
$T = 100$, Negligible Group Auditor Priors

	Component Materiality T_i						Relative Total Variable Cost $RTVC$						Achieved Group Assurance AGA					
N	GUAM	MACM	SQRT	PROP	HALF	FULL	GUAM	MACM	SQRT	PROP	HALF	FULL	GUAM	MACM	SQRT	PROP	HALF	FULL
2	63	75	71	50	50	100	1.58	1.33	1.41	2.00	2.00	1.00	95%	91%	92%	98%	98%	80%
3	48	67	58	33	50	100	2.10	1.50	1.73	3.00	2.00	1.00	95%	83%	89%	99%	94%	58%
4	39	50	50	25	50	100	2.59	2.00	2.00	4.00	2.00	1.00	95%	85%	85%	> 99%	85%	35%
5	33	50	45	20	50	100	3.06	2.00	2.24	5.00	2.00	1.00	95%	71%	80%	> 99%	71%	18%
6	28	42	41	17	50	100	3.51	2.40	2.45	6.00	2.00	1.00	95%	72%	74%	> 99%	55%	8%
7	25	43	38	14	50	100	3.95	2.33	2.65	7.00	2.00	1.00	95%	55%	68%	> 99%	39%	3%
8	23	38	35	13	50	100	4.39	2.67	2.83	8.00	2.00	1.00	95%	55%	61%	> 99%	25%	1%
9	21	33	33	11	50	100	4.82	3.00	3.00	9.00	2.00	1.00	95%	54%	54%	> 99%	15%	< 1%
10	19	35	32	10	50	100	5.24	2.86	3.16	10.00	2.00	1.00	95%	35%	47%	> 99%	8%	< 1%

The formulas for T_i and $RTVC$ are provided in Table 1 for each of the methods. AGA is defined in Formula (21). Example: for $N = 3$, the MACM formulas yield $T_i = mT/N = 2.0 \times 100/3 = 67$, $RTVC = N/m = 3/2.0 = 1.50$, and $AGA = G(T;N,T_i/3.0) = G(100;3,67/3.0) = 83\%$.

FIGURE 3
Comparison of T_i, *RTVC*, and *AGA* for Various Methods with $N = 2$ to 10 Equally Sized, Equally Weighted Components



These charts compare alternative component materiality methods by three criteria: Component Materiality (T_i), Relative Total Variable Cost (*RTVC*), and Achieved Group Assurance (*AGA*) assuming $T = 100$ and negligible priors.

Minimizing Audit Cost for Components of Unequal Sizes

For GUAM, *any* combination of positive weights summing to one will result in $AGA \approx 95$ percent and, therefore, the group auditor is free to choose weights that also achieve secondary objectives. If the objective is to minimize component audit costs, then components should be weighted approximately in proportion to the square root of their size:

$$w_i = \frac{\sqrt{Y_i}}{\sum \sqrt{Y_j}}. \tag{24}$$

This result is derived in Appendix B as the solution to a classic constrained optimization problem. If components are equally sized, then Equation (24) simplifies to $w_i = 1/N$.

In Table 3 we compare *AGA* and *RTVC* results across component materiality methods for an illustrative three-component group in which component sizes vary by a factor of 9 and where Equation (24) is used to weight components applying GUAM. GUAM achieves $AGA = 95$ percent and $RTVC = 1.87$, while for the other methods only PROP achieves $AGA \geq 95$ percent, but at higher *RTVC*. The other methods fail to achieve the target $AGA \geq 95$ percent, and HALF is both ineffective and inefficient relative to GUAM.

IV. GUAM FOR PRACTICAL GROUP CONTEXTS AND STRUCTURES

In this section we adapt GUAM to optimize component materiality for various practical contexts using an *efficient materiality frontier* consisting of all T_i combinations for which target *AGA* is achieved but not exceeded. We conclude by showing how group structure may allow the group auditor to form a prior assurance profile over a subgroup of components, thereby facilitating substantial reduction of component audit cost.

TABLE 3

Comparative T_i, $RTVC$, and AGA for Various Component Materiality Methods: Unequally Sized Components
$T = 100$, Negligible Group Auditor Priors

	Size(Y_i)	GUAM	MACM	SQRT	PROP	HALF	FULL
Component Materiality(T_i)							
Component 1	9,000	63	100	80	64	50	100
Component 2	4,000	48	67	53	29	50	100
Component 3	1,000	28	33	27	7	50	100
	14,000	139	200	160	100	150	300
RTVC		1.87	1.29	1.60	3.00	2.00	1.00
AGA		95%	82%	91%	98%	94%	58%

This table illustrates T_i, $RTVC$, and AGA for unequally sized components under the various component materiality methods described in this paper. See Table 1 for T_i and $RTVC$ calculations and Formula (21) for AGA. The GUAM calculations are performed using square-root weighting Formula (24). For example, for Component 1, $w_1 = \sqrt{9000}/(\sqrt{9000} + \sqrt{4000} + \sqrt{1000}) = 0.5$, and from Equation (18), $T_1 = TR_{.95}/Q_{.95}(1/w_1) = 100 \times 3.0/4.74 = 63$.

Optimizing T_i along the "Efficient Materiality Frontier"

In group audit practice, component materiality may be constrained by factors such as statutory audit requirements, specific user expectations, or jurisdiction-specific financial reporting framework requirements. Also, VCU_i may vary significantly between components or a component might be easy to audit "one-hundred percent" and thus effectively reducing the probability of that component's misstatement to zero. In such circumstances, various combinations of component materiality amounts may be optimal. For any combination $(w_1,\ldots,w_N)$ of component weights such that $w_1 + \cdots + w_N = 1$ there is a corresponding combination $(T_1,\ldots,T_N)$ of GUAM-based component materiality values that together yield $AGA = 95$ percent. The T_i combinations $(T_1,\ldots,T_N)$ constitute an *efficient materiality frontier* (EMF).

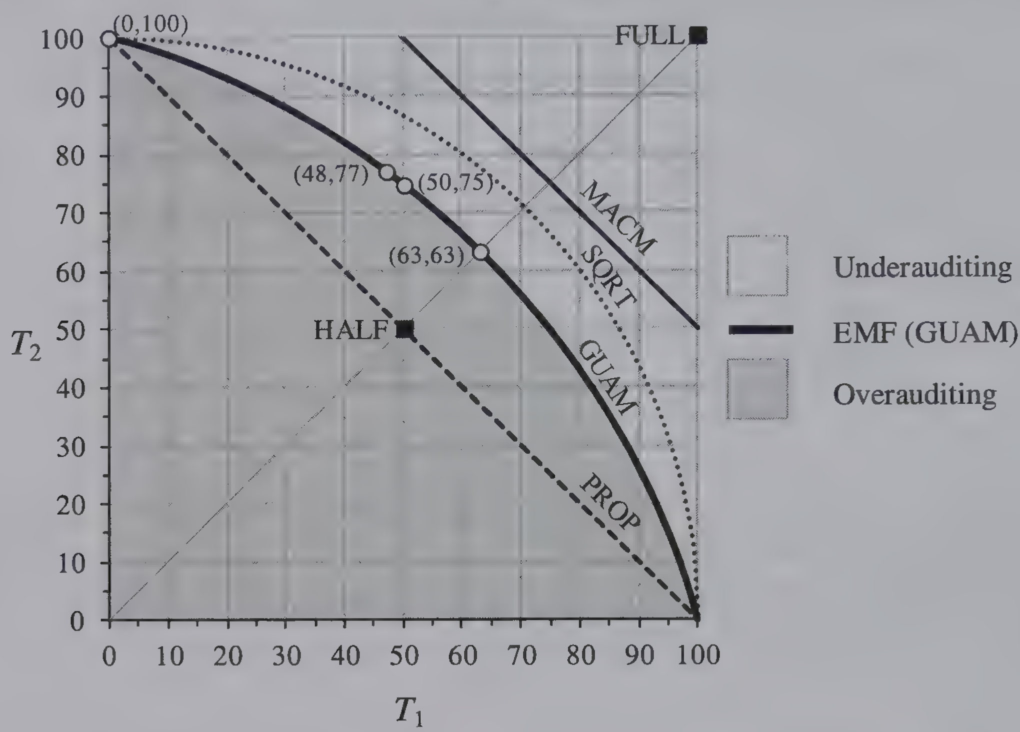
For two components, the (T_1,T_2) EMF can be represented by a continuous curve such as that depicted in Figure 4 for $T = 100$ and negligible component priors. The EMF partitions all possible component materiality combinations (T_1,T_2) into those above the EMF curve (ineffective because $AGA < 95$ percent), those below the EMF (effective but inefficient because AGA $>$ 95 percent), and those on the EMF (effective and efficient with AGA $=$ 95 percent). Each combination on the EMF may be optimal for specific contexts depending upon costs, constraints, and other contextual factors and, as elaborated below, the assignment of appropriate weights allows the determination of optimal component materiality combinations on the EMF.[19]

Curves plotting (T_1,T_2) combinations for MACM, SQRT, and PROP and the fixed points corresponding to HALF and FULL are shown for comparison in Figure 4. Note that *all* combinations using SQRT and MACM as well as the FULL combination are above the EMF and, thus, ineffective. Also, all combinations using PROP are below the EMF and, thus,

[19] The effect of including one or more component priors would be to shift the EMF curve outward, away from the origin, reflecting greater component materiality amounts.

FIGURE 4

Efficient Materiality Frontier: Two Components, Negligible Group Auditor Priors, $T = 100$



The two-component Efficient Materiality Frontier (EMF), labeled GUAM, is the set of component materiality pairs (T_1, T_2) that achieve $AGA = 95\%$. Curves for the PROP, SQRT, and MACM methods are shown as well as the fixed points for HALF $= (50, 50)$ and FULL $= (100, 100)$. Pairs above the EMF result in under-auditing, and those below in over-auditing. Points on the EMF illustrating optimality under various conditions are indicated:
$(T_1,T_2) = (63,63)$—Components are indistinguishable;
$(T_1,T_2) = (48,77)$—Component 1 is a quarter the size of Component 2;
$(T_1,T_2) = (48,77)$—Component 1 variable audit cost is a quarter that of Component 2;
$(T_1,T_2) = (50,75)$—Component 1 materiality is constrained to $T_1 = 50$; and
$(T_1,T_2) = (0,100)$—Component 1 is audited "100 percent."

effective, but inefficient relative to GUAM. Finally, HALF, like PROP, is effective, but inefficient for $N = 2$.

To illustrate the construction of optimal (T_1,T_2) combinations to meet a secondary objective (beyond group AGA = 95 percent), Figure 4 plots the optimal combination for the contexts described as follows:

Equal Size: $Y_1 = Y_2$. Because there is nothing else to distinguish them, $(w_1,w_2) = (1/2,1/2)$ and, per Equation (19), $(T_1,T_2) = (63,63)$. Per Equation (23), minimum $RTVC = 100/63 = 1.58$.

Different Sizes: $Y_1 = 0.25Y_2$. Square-root weighting per Equation (24) is applied to minimize $RTVC$ to yield:

$$w_1 = \frac{\sqrt{Y_1}}{\sqrt{Y_1} + \sqrt{Y_2}} = \frac{\sqrt{0.25Y_2}}{\sqrt{0.25Y_2} + \sqrt{Y_2}} = \frac{\sqrt{0.25}}{\sqrt{0.25} + \sqrt{1}} = \frac{1}{3},$$

$(w_1,w_2) = (1/3,2/3)$ and, per Equation (18), $(T_1,T_2) = (48,77)$. Per Equation (23), minimum $RTVC = (0.25/48 + 1/77)/(1.25/100) = 1.46$.

Different Costs: $(VCU_1,VCU_2) = (0.25,1.00)$. By replacing Y_i with $VCU_i \times Y_i$ in (24) and assuming $Y_1 = Y_2$:

$$w_1 = \frac{\sqrt{VCU_1 Y_1}}{\sqrt{VCU_1 Y_1} + \sqrt{VCU_2 Y_2}} = \frac{\sqrt{0.25 Y_1}}{\sqrt{0.25 Y_1} + \sqrt{1.00 Y_2}} = \frac{\sqrt{0.25}}{\sqrt{0.25} + \sqrt{1.00}} = \frac{1}{3},$$

$(w_1,w_2) = (1/3,2/3)$ and, per Equation (18), $(T_1,T_2) = (48,77)$ and minimum $RTVC = 1.46$ as in the previous example.

Component Materiality Constraint: $T_1 = 50$ by mandate. From Equation (18), mandating $T_1 = 50$ implies that the weight w_1 is the solution to:

$$Q_{0.95}(1/w_1) = \frac{R_{0.95}T}{T_1} = \frac{3.0 \times 100}{50} = 6.0.$$

Solving for w_1 yields 0.36 and setting $w_2 = 1 - w_1$, gives $w_2 = 0.64$. From Equation (18), $(T_1,T_2) = (50,75)$ and, per Equation (23), minimum $RTVC = (1/50 + 1/75)/(2/100) = 1.67$. Thus, $T_1 = 50$ allows T_2 to increase to 75. This approach to accommodating constrained component materiality can be generalized to multiple component constraints.[20]

One-Hundred Percent Examination: $X_1 = 0$ with 100 percent assurance. Component 1 is audited enough to reduce its audit risk to zero so that possible misstatements are limited to Component 2 and $T_2 = T$ is sufficient for group purposes and the optimal EMF component materiality pair is $(T_1,T_2) = (0,100)$.[21]

Exploiting Group Structure and "Clusters"

To this point we have assumed that groups are comprised of separately managed stand-alone components so that the group auditor *cannot* form a conclusion on the group financial statements by means other than aggregating individual component audit conclusions. In practice, many groups assemble subgroups of components with similar business activities and these components share subgroup-level resources and oversight that provide a degree of cohesiveness.[22] In U.S. regulations governing audits of states, local governments, and non-profit organizations (Office of Management and Budget [OMB] 2003), the term "cluster" is used for "a grouping of closely related programs that share common compliance requirements." We refer to a subgroup operated under common group-level controls as a *cluster*.

Where a cluster exists, the group auditor may be able to treat it as a "virtual single component" for group audit-planning purposes by forming a prior assurance profile for the cluster as a whole.[23]

---

[20] Assume that $T_1,\ldots,T_k$ are unconstrained but $T_{k+1},\ldots,T_N$ are constrained to amounts less than those required for group audit purposes. For constrained components, implied w_i is the solution to Equation (18), $Q_{0.95}(1/w_i) = R_{0.95}T/T_i$ and the sum of their weights is $S = \sum_{i=k+1}^{N} w_i$. For unconstrained components $1,\ldots,k$, we modify Equation (24) so that $w_i = (\sqrt{Y_i}/\sum_{j=1}^{k}\sqrt{Y_j})(1-S)$ and use Equation (18) to compute T_i. Note that $\sum_{i=1}^{N} w_i = 1$ as required.

[21] Because Component 1 is audited 100 percent, the notion of *VCU* does not apply to Component 1 and we do not compute *RTVC*. The optimum, $(T_1,T_2) = (0,100)$, is also approached as Y_1 or VCU_1 approach zero.

[22] As an example, a group's retail operation might be comprised of dozens of separately operated outlets that share subgroup-level supply chain management, advertising, incentives, and monitoring.

[23] ISA 600, paragraph A3, also recognizes that the group auditor may "identify components at certain levels of aggregation rather than individually."

Such a cluster prior might be justified by the group auditor's evaluation of cluster-wide risks, by tests of the design and operating effectiveness of cluster-wide controls and/or by the results of cluster-level analytical procedures. With such a prior, the group auditor can establish a target cluster posterior by applying Equation (14) as if for a single component, and then applying Equations (15) through (20) to determine *cluster materiality* for the cluster as a whole.

To derive component materiality for cluster members, suppose $g(\sum_L x_j; 1, \beta)$ is the prior assurance profile for a cluster of L components *as a whole*, where the random variable $\sum_L X_j$ represents prior total misstatement for the cluster. Assuming as before that the X_j are stochastically independent, $g(\sum_L x_j; 1, \beta)$ is the aggregate of the individual prior distributions of $X_1, \ldots, X_L$.[24] The individual component priors comprising the aggregate are $g(x_j;\alpha_j,\beta)$, where the *fractional* α_j values that comprise the cluster prior are the component's relative size (i.e., $\alpha_j = Y_j / \sum_L Y_l$) and thus $\sum_L \alpha_j = 1$ (Stewart 2012). This disaggregation from an overall cluster prior into fractional component priors can be expressed schematically as:

$$\overset{\text{Cluster Prior}}{g(x_1 + x_2; 1, \beta)} \overset{\text{Disaggregate}}{\longrightarrow} \overset{\text{Component Priors}}{\begin{bmatrix} g(x_1; 0.5, \beta) \\ g(x_2; 0.5, \beta) \end{bmatrix}}.$$

As to the cluster posterior and likelihood function, if the target cluster posterior is $g(\sum_L x_j; 1, \beta')$, then target component posteriors are $g(x_j;\alpha_j,\beta')$ with $\sum_L \alpha_j = 1$. The likelihood function required to update the cluster prior to the cluster posterior is $g(\sum_L x_j; 1, B)$, where B is defined by Equation (10). Similarly, an exponential likelihood function with parameter B is also required to update the component prior $g(x_j;\alpha_j,\beta)$ to the component posterior $g(x_j;\alpha_j,\beta')$. Because cluster and component materiality are each the 95th percentile of the relevant likelihood function, and cluster and component likelihoods are equal, component materiality equals cluster materiality for each component in the cluster.

Figure 5 illustrates the process for a cluster of two equally sized components. The group auditor must begin by establishing an exponential prior on the cluster, which in this case is $g(x_1 + x_2;1, 100/0.7)$, and then the target cluster posterior $g(x_1 + x_2;1, 100/3.0)$, so the cluster likelihood function is $g(x_1 + x_2;1,100/2.3)$. The cluster prior disaggregates into fractional component priors,[25] schematically:

$$\overset{\text{Cluster Prior}}{g(x_1 + x_2; 1, 100/0.7)} \overset{\text{Disaggregate}}{\longrightarrow} \overset{\text{Component Priors}}{\begin{bmatrix} g(x_1; 0.5, 100/0.7) \\ g(x_2; 0.5, 100/0.7) \end{bmatrix}}.$$

The likelihood functions are the same at the cluster and component levels so cluster and component materialities are the same. Assuming that the audit goes as planned, the component posteriors aggregate to achieve the desired cluster posterior, schematically:

$$\overset{\text{Component Posteriors}}{\begin{bmatrix} g(x_1; 0.5, 100/3.0) \\ g(x_2; 0.5, 100/3.0) \end{bmatrix}} \overset{\text{Aggregate}}{\longrightarrow} \overset{\text{Cluster Posterior}}{g(x_1 + x_2; 1, 100/3.0)}.$$

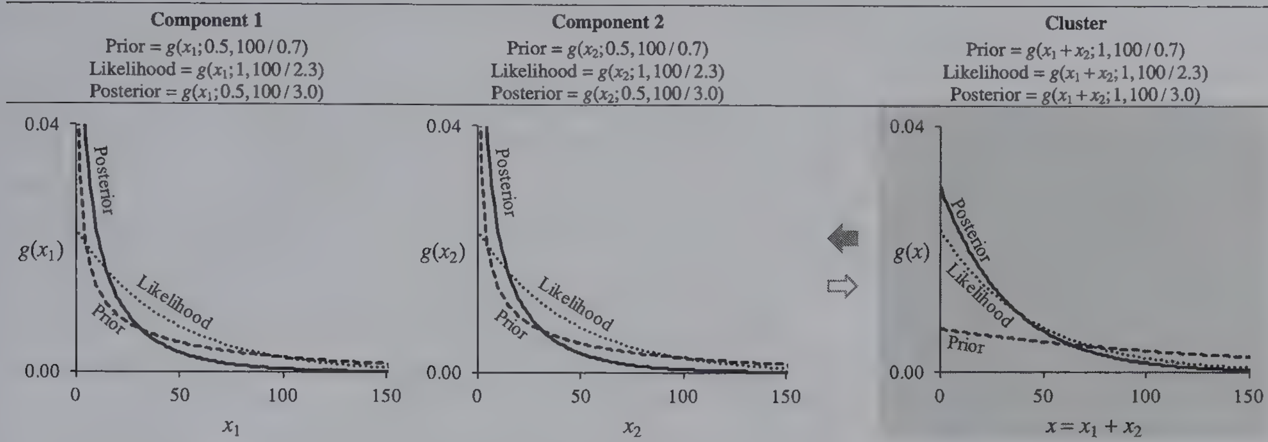
In other words, the two fractional component posteriors ($\alpha_1', \alpha_2' = 0.5$) aggregate to a cluster exponential posterior ($\alpha' = \alpha_1' + \alpha_2' = 1$).

---

[24] If cluster components operate under commonalities such as cluster-level incentives and performance pressures, then the independence assumption may not apply.

[25] Note that in Figure 5 the fractional component priors disaggregate the $\alpha = 1$ exponential cluster prior.

FIGURE 5
Disaggregation of the Cluster Prior into Component Priors and Aggregation of Component Posteriors into the Cluster Posterior



← Plan (disaggregation): The group auditor establishes an exponential prior $g(x_1 + x_2;1,100/0.7)$ on a cluster of two equally sized components and a target cluster posterior $g(x_1 + x_2;1,100/3.0)$, which disaggregate into component priors $g(x_i;0.5,100/0.7)$ and target component posteriors $g(x_i;0.5,100/3.0)$, $i = 1, 2$, respectively. The likelihood functions are the same at the cluster and component levels, and, therefore, cluster and component materialities are the same.
→ Conclude (aggregation): If the audits go as planned, then the component posteriors aggregate to achieve the desired cluster posterior.

Illustration of the Effects of Cluster-Level Prior Assurance

We close with an extended example that illustrates the dramatic reduction of component audit cost possible when the group auditor can form prior cluster-level assurance. Figure 6 presents a group comprised of ten equally sized units across three business groupings: six Retail components, three Manufacturing components, and one Real Estate component. In Panel A all of the components are stand-alone and there is negligible prior assurance for each component. Each component materiality is $T_i = 19$ (see Table 2) and $RTVC = 5.24$ with the Retail components comprising 3.14 of the total.

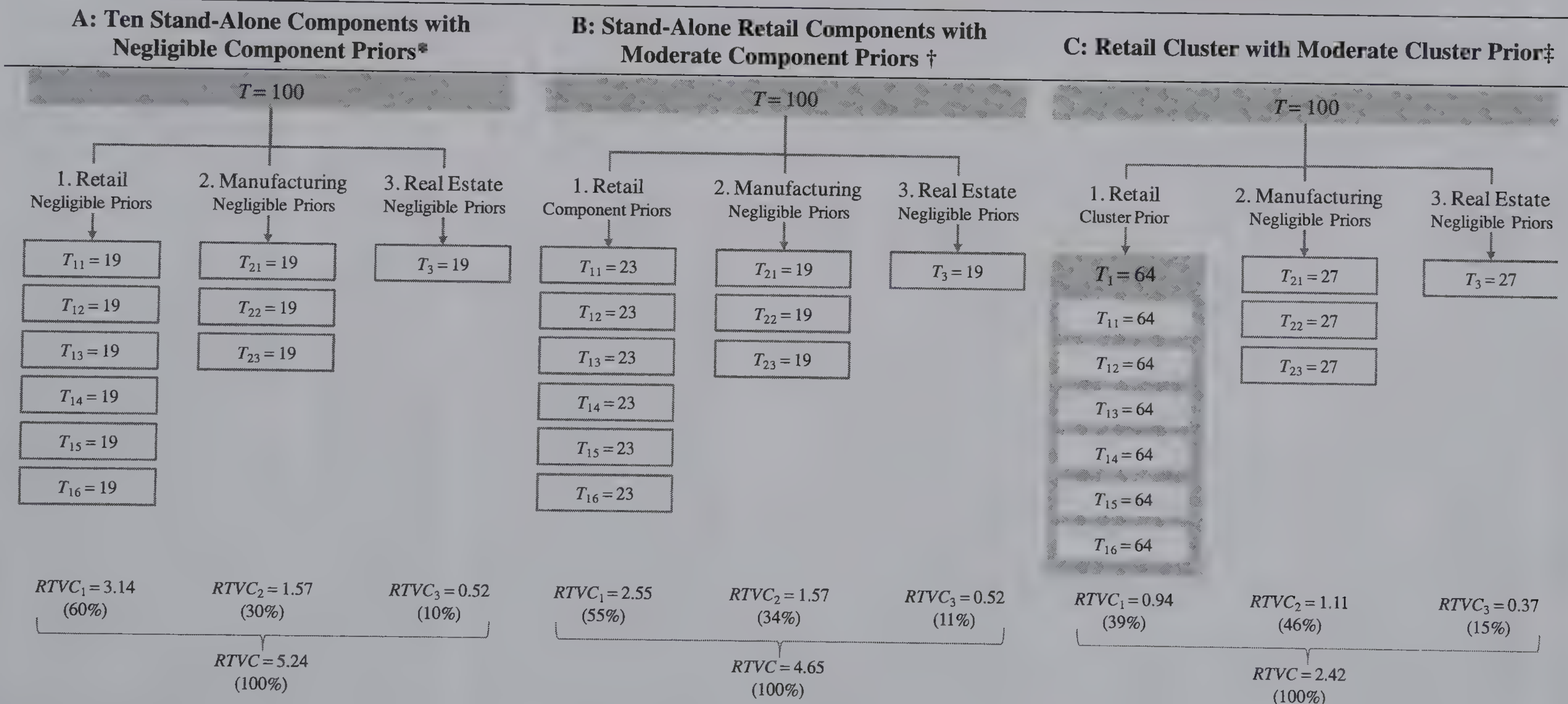
In Figure 6, Panel B the group auditor has evaluated group-level controls for each Retail component individually and determined that there is moderate (RMM_{1j} = 50 percent) prior assurance for each. Component materiality for each Retail component is 23, which reduces *RTVC* across the six Retail components to 2.55, while the other four components remain the same and group *RTVC* drops to 4.65.

In Figure 6, Panel C the group auditor is able to form a cluster prior for the Retail components and assesses a moderate profile ($RMM_1 = 50$ percent) for the cluster. Cluster materiality is 64 and, because there is a basis for treating the six units as a virtual single component, each of the six Retail components is audited using $T_{1j} = 64$ and *RTVC* drops to 0.94 or 30 percent of the *RTVC* of Retail units in Panel A. Also, even though nothing has changed for the three Manufacturing and one Real Estate units, component materialities are now 27 and *RTVC* drops from 1.57 to 1.11, and from 0.52 to 0.37 for the three Manufacturing units and the Real Estate unit, respectively.

For the group as a whole, the presence of the Retail cluster and its moderate prior allows a reduction in *RTVC* of more than 50 percent with reductions in every component while maintaining

FIGURE 6

T_i and *RTVC* for Ten Equally Sized Components in Three Industry Segments under Three Prior Assurance Scenarios



* Ten stand-alone components with negligible group auditor component priors throughout.

† Stand-alone retail components have moderate priors, $g(x_{1j};1,23/0.7)$, $j = 1,\ldots,6$ (equivalent to $RMM_{1j} = 50$ percent at the component level).

‡ Retail components comprise a cluster with a moderate cluster prior $g(x_1;1,64/0.7)$ (equivalent to $RMM_1 = 50$ percent at the cluster level).

AGA at 95 percent. This significant change occurs because the group in Panel C effectively consists of five stand-alone components: one large virtual single component (the cluster) plus four individual stand-alone components compared to the ten stand-alone components in Panels A and B.

This illustration shows the sizeable reduction in component audit cost that is possible from the group auditor's reliance on cluster-level prior assurance. The example does not address the development of reliable prior assurance for a cluster as a whole, which is likely to be costly. Such a prior would require the existence, identification, and analysis of a group's business activities that have substantial cohesiveness, exploitable similarities, and common controls as well as tests of the design and operating effectiveness of cluster-level controls. However, the potential for component audit cost savings via clusters highlights the importance of conceptual models, archival analyses of extant audit practices, and behavioral studies to support reliable auditing procedures for assessing cluster and other group-level controls and risks.

V. CONCLUSION AND LIMITATIONS

This paper addresses the fundamental auditing problem of how to plan component audits so that conclusions about separately audited component information can be aggregated to achieve the group auditor's assurance objective at reasonable cost. We do so by generalizing the single component audit risk model to derive a unified assurance and materiality model (GUAM). Our results provide a comprehensive first pass at modeling group audits applying ISA 600 and suggest conceptual and empirical solutions to practice problems.

Using the gamma family of probability distributions and Bayesian probability concepts, we derive component materiality combinations that, assuming audits go as planned, achieve group assurance targets as required by auditing standards while minimizing total component audit cost. The method incorporates the group auditor's prior assurance on components and statutory audit requirements for components as well as various configurations of group-level controls, group structures, and components of disparate sizes. GUAM may form the basis for other conceptual models of audits as well as for research focused on elements of group audits.

We show that alternative methods used in group audit practice result in component materialities that vary widely as to achieved assurance and cost, depending on circumstances. These alternative methods may be ineffective in meeting professional standards, inefficient in the use of resources, or both ineffective and inefficient. We believe that these differences in results suggest the need for guidance more definitive than that provided by current auditing standards and that the GUAM approach can be a useful metric for evaluating the efficacy of other, less comprehensive, methods for determining component materiality in practice.

At least three limitations deserve exploration. First, because little is known about the effectiveness and efficiency of extant and emerging group audit practices, there is a need for comparative archival and behavioral research to support best practices and to facilitate standards setting for planning group audits under practical conditions. Second, we simply assume specific values for group-level component and cluster priors. In practice, group auditors must form and be able to defend their particular priors for components and clusters. Research is needed regarding the conceptual underpinnings, design, and operating effectiveness of group-level controls and their impact on these priors. Finally, we assume that the random variables representing the group auditor's assessments of potential component misstatement are stochastically independent. In some practical group contexts, this assumption is unlikely to be satisfied and suggests a need for both empirical and analytical research into addressing the causes, forms, and consequences of non-independence.

REFERENCES

Akresh, A. 2010. A risk model to opine on internal control. *Accounting Horizons*. 24 (1): 56–78.

Akresh, A., J. Loebbecke, and W. Scott. 1988. Audit approaches and techniques. In *Research Opportunities in Auditing: The Second Decade*, edited by K. Abdel-Khalik and I. Solomon. Sarasota, FL: American Accounting Association.

American Institute of Certified Public Accountants (AICPA). 2006. *Audit Risk and Materiality in Conducting an Audit*. AU Section 312. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). 2008. *Audit Sampling*. Audit Guide. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). 2011. *Special Considerations—Audits of Group Financial Statements (Including the Work of Component Auditors)*. AU-C Section 600. New York, NY: AICPA.

Bailey, A. 1981. *Statistical Auditing: Review, Concepts, and Problems*. New York, NY: Harcourt Brace Jovanovich, Inc.

Blokdijk, H., F. Drieënhuizen, and P. Wallage. 1995. Deventer: Kluwer Bedrijfswetenschappen. Kluwer/ Limperg Instituut-reeks, Vol. 4. Amsterdam, The Netherlands: Limperg Instituut.

Boritz, E., P. Zhang, and S. Aldersley. 1993. On combining evidence from subpopulations into a composite conclusion. *Contemporary Accounting Research* 10 (Fall): 227–245.

Cooper, L., and D. Steinberg. 1970. *Introduction to Methods of Optimization*. Philadelphia, PA: W. B. Saunders Company.

DeGroot, M., and M. Schervish. 2002. *Probability and Statistics*. 3rd ed. Boston, MA: Addison-Wesley.

Dutta, S., and L. Graham. 1998. Considering multiple materialities for account combinations in audit planning and evaluation: A cost efficient approach. *Journal of Accounting, Auditing and Finance* (Spring): 151-171.

Felix, W. 1974. A decision theory view of auditing. *Contemporary Accounting Problems: Proceedings of the 1974 Arthur Andersen/University of Kansas Symposium on Auditing Problems*, 63–71. Lawrence, KS: University of Kansas Printing Service.

Glover, S., D. Prawitt, J. Liljegren, and W. Messier. 2008a. Component materiality for group audits. *Journal of Accountancy*. (December). Available at: http://www.journalofaccountancy.com/Issues/2008/Dec/ComponentMaterialityforGroupAudits.htm

Glover, S., D. Prawitt, J. Liljegren, and W. Messier. 2008b. Component materiality for group audits: Probabilistic model. *Journal of Accountancy*. Available at: http://www.journalofaccountancy.com/Web/ProbabilisticModel.htm

International Auditing and Assurance Standards Board (IAASB). 2007. *Special Considerations—Audits of Group Financial Statements (Including the Work of Component Auditors)*. International Standard on Auditing 600. New York, NY: International Federation of Accountants.

International Auditing and Assurance Standards Board (IAASB). 2008a. *Overall Objectives of the Independent Auditor and the Conduct of an Audit in Accordance with International Standards on Auditing*. International Standard on Auditing 200. New York, NY: International Federation of Accountants.

International Auditing and Assurance Standards Board (IAASB). 2008b. *Materiality in Planning and Performing an Audit*. International Standard on Auditing 320. New York, NY: International Federation of Accountants.

Jaynes, E. 1957. Information theory and statistical mechanics. *The Physical Review* 106 (4): 620–630. Available at: http://bayes.wustl.edu/etj/articles/theory.1.pdf

Johnson, N., S. Kotz, and N. Balakrishnan. 1994. *Continuous Univariate Distributions* Vol. 1, 2nd edition. New York, NY: John Wiley & Sons.

Jones, P. 1999. *Statistical Sampling and Risk Analysis in Auditing*. Aldershot, U.K.: Gower.

Kinney, W. 1975. Decision theory aspects of internal control system design/compliance and substantive tests. *Journal of Accounting Research* (Supplement): 14–29.

Kinney, W. 1979. Integrating audit tests: Regression analysis and partitioned dollar-unit sampling. *Journal of Accounting Research* 17 (Fall): 456-475.

Kinney, W. 1984. Discussant's response to Leslie (1984): An analysis of the audit framework focusing on inherent risk and the role of statistical sampling in compliance testing. In *Proceedings of the 1984 Touche Ross University of Kansas Symposium on Auditing Problems*, edited by H. Stettler, and N. Ford, 126–132. Lawrence, KS: School of Business, University of Kansas.

Kinney, W. 1989. Achieved audit risk and the audit outcome space. *Auditing: A Journal of Practice & Theory* 8 (Supplement): 67–84.

Kinney, W. 1993. Audit risk at the financial statement level: The joint misstatement problem. Working paper (revised 2003), The University of Texas at Austin.

Leslie, D. 1984. An analysis of the audit framework focusing on inherent risk and the role of statistical sampling in compliance testing. In *Proceedings of the 1984 Touche Ross University of Kansas Symposium on Auditing Problems*, edited by H. Stettler, and N. Ford, 89–125. Lawrence, KS: School of Business, University of Kansas.

Leslie, D. 1985. *Materiality: The Concept and Its Application to Auditing*, CICA Research Study. Toronto, Canada: The Canadian Institute of Chartered Accountants.

Leslie, D., A. Teitlebaum,and R. Anderson. 1979. *Dollar-Unit Sampling*. Toronto, Canada: Copp Clark Pitman.

Messier, W., N. Martinov-Bennie, and A. Eilifsen. 2005. A review and integration of empirical research on materiality: Two decades later. *Auditing: A Journal of Practice & Theory* 24 (2): 153–187.

Messier, W., S. Glover, and D. Prawitt. 2010. *Auditing & Assurance Services: A Systematic Approach*. 7th edition. New York, NY: McGraw-Hill Irwin.

Office of Management and Budget (OMB). 2003. Circular A-133. *Audits of States, Local Governments, and Non-Profit Organizations*. Circular A-133. Washington, DC: Government Printing Office.

O'Hagan, A., and J. Forster. 2004. *Kendall's Advanced Theory of Statistics, Vol. 2B, Bayesian Inference*. 2nd edition. London, U.K.: Arnold.

Public Company Accounting Oversight Board (PCAOB). 2010. *Consideration of Materiality in Planning and Performing an Audit*. Auditing Standard No. 11. Washington, DC: PCAOB.

Raiffa, H., and R. Schlaifer. 2000. *Applied Statistical Decision Theory*. New York, NY: Wiley Classics Library.

Schilder, A. 1995. Reduction of uncertainty: The core function of auditors. In *Chance and Uncertainty*, edited by H. Capel, J. Cramer, O. Estévez-Uscanga, C. Klaassen, and G. Mellenberg, 195–213. Amsterdam, The Netherlands: Vossiuspers AUP.

Smith, T. 1976. *Statistical Sampling for Accountants*. London, U.K.: Haymarket.

Steele, A. 1992. *Audit Risk and Audit Evidence: The Bayesian Approach to Statistical Auditing*. San Diego, CA: Academic Press.

Stewart, T. 2008. *Audit Sampling*. Technical notes on the AICPA audit guide, new edition, May 2008. New York, NY: AICPA. Available at: http://www.aicpa.org/Publications/AccountingAuditing/KeyTopics/Pages/AuditSampling.aspx

Stewart, T. 2012. *A Bayesian Audit Assurance Model with Application to the Component Materiality Problem in Group Audits*. Ph.D. dissertation, VU University Amsterdam.

Stewart, T., L. Strijbosch, H. Moors, and P. van Batenburg. 2007. A simple approximation to the convolution of gamma distributions. Discussion Paper, No. 2007-70, Tilburg University. Available at: http://arno.uvt.nl/show.cgi?fid=63976

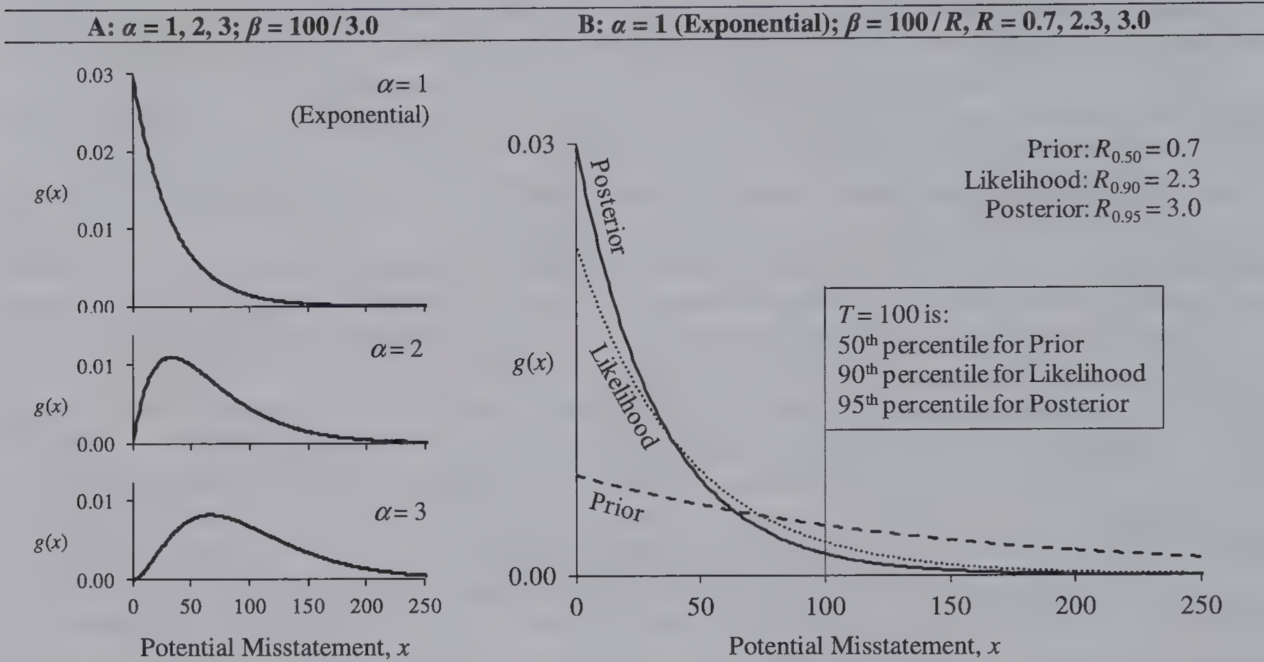
Stringer, K., and T. Stewart. 1996. *Statistical Techniques for Analytical Review in Auditing*. 2nd edition. New York, NY: John Wiley & Sons, Inc.

Turner, J. 1997. The impact of materiality decisions on financial ratios: A computer simulation. *Journal of Accounting, Auditing and Finance* 12 (Spring): 125–147.

Van Batenburg, P., A. O'Hagan, and R. Veenstra. 1994. Bayesian discovery sampling in financial auditing: A hierarchical prior model for substantive test sample sizes. *The Statistician* 43 (1): 99–110.

Zuber, G., R. Elliott, W. Kinney, and J. Leisenring. 1983. Using materiality in audit planning. *Journal of Accountancy* (March).

EXHIBIT 1
Illustrative Gamma Distributions



Panel A depicts three gamma distributions $g(x;\alpha,\beta)$ with shape parameters $\alpha = 1$, 2, and 3 and scale parameter $\beta = 100/3.0$, distributions that might apply to an audit conducted using $T = 100$. Panel B illustrates how the exponential distribution $g(x;1,100/3.0)$ might have arisen from a prior distribution $g(x;1,100/0.7)$ based on $RMM = 50$ percent and a likelihood function $g(x;1,100/2.3)$ induced by further audit procedures.

APPENDIX A

GAMMA DISTRIBUTION: BACKGROUND CONCEPTS

This appendix defines the gamma probability density, cumulative distribution, and percentile functions; describes important properties of the gamma distribution and the various shapes that arise in the GUAM model; and explains Bayes' rule in the context of the gamma distribution.

Description and Application

The gamma distribution is defined by the *probability density* function:

$$g(x;\ \alpha,\beta) = \frac{(x/\beta)^{\alpha-1} e^{-x/\beta}}{\beta\,\Gamma(\alpha)} \quad \text{for} \quad x > 0 \quad \text{and} \quad \alpha,\beta > 0, \tag{25}$$

where x is the value of a random variable X and $\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt$ (Johnson et al. 1994, 337).[26] Gamma distributions with $\alpha = 1$, 2, and 3 and the same scale parameter $\beta = 100/3.0$ are depicted in Exhibit 1, Panel A and three exponential ($\alpha = 1$) distributions are depicted in Panel B.

Exponential distributions arise in a component audit when there is no evidence of misstatement and thus projected misstatement is zero. Distributions with $\alpha > 1$ arise when there is evidence of

[26] The *gamma function* Γ has the recursive property, $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$, and $\Gamma(1) = 1$, which for positive integer values of α relates it to the factorial function through $\Gamma(\alpha) = (\alpha - 1)(\alpha - 2)\cdots 2\cdot 1 = (\alpha - 1)!$

misstatement and the component distribution peaks at the most probable total misstatement. They also arise in group audits when multiple component distributions are aggregated to form the distribution for a group—even if there is no evidence of misstatement in any component. Distributions with $\alpha < 1$ are more skewed and peaked at zero than the exponential distribution and arise with "clusters" as discussed in Section IV and illustrated in Figure 5.

The *mode* of the gamma distribution is Mode $(X) = (\alpha - 1)\beta$ for $\alpha \geq 1$ and zero for $\alpha < 1$; its *expected value* (or *mean*) is $E(X) = \alpha\beta$; its *variance* is $\mathrm{Var}(X) = \alpha\beta^2$; and its *standard deviation* is $\sigma(X) = \sqrt{\alpha}\beta$. The parameters α and β expressed in terms of the mean and variance are:

$$\alpha = \frac{[E(X)]^2}{\mathrm{Var}(X)} \quad \text{and} \quad \beta = \frac{\mathrm{Var}(X)}{E(X)}.$$

As can be seen in Exhibit 1, the gamma distribution becomes more bell-shaped as the parameter α increases: it tends to the normal distribution as α tends to infinity (Johnson et al. 1994, 340), so for large α, $g(x; \alpha, \beta) \approx Normal(x; \alpha\beta, \sqrt{\alpha}\beta)$.

Cumulative Distribution

The *cumulative distribution* function:

$$G(x; \alpha, \beta) = \int_0^x g(t; \alpha, \beta)\, dt$$

gives the probability that the value of X does not exceed x, that is, $G(x;\alpha,\beta) = \Pr(X \leq x)$. Graphically, it is the area under $g(x;\alpha,\beta)$ between 0 and x.

Percentile Function

The *percentile* function $Q(P;\alpha,\beta)$ is the inverse of $G(x;\alpha,\beta)$. If $G(x;\alpha,\beta) = P < 1$, then $x = Q(P;\alpha,\beta)$ is the 100Pth percentile of X. Because β is a scale parameter, $Q(P;\alpha,\beta)$ is equal to $\beta Q(P;\alpha,1)$, which for conciseness we write as $\beta Q_P(\alpha)$.[27] For example, $\beta Q_{0.95}(\alpha)$ is the 95th percentile of a random variable X distributed as $g(x;\alpha,\beta)$.

The Exponential Distribution

In the case of exponential ($\alpha = 1$) distributions, Mode$(X) = 0$, $E(X) = \beta$, and $\mathrm{Var}(X) = \beta^2$ and the three probability functions discussed above have simple forms that are amenable to direct calculation. The probability density function is:

$$g(x; 1, \beta) = \frac{1}{\beta} e^{-x/\beta}, \tag{26}$$

the cumulative distribution function is:

$$G(x; 1, \beta) = \int_0^x \frac{1}{\beta} e^{-t/\beta} dt = 1 - e^{-x/\beta},$$

and the percentile function, obtained by solving $G(x;1,\beta) = P$ for x, is:

[27] In words, the 100Pth percentile of a gamma distribution with scale β is equal to β times the 100Pth percentile of a gamma distribution with scale 1. Also $G(x;\alpha,\beta) = G(x/\beta;\alpha,1)$.

$$x = -\beta \ln(1 - P)$$

which is denoted $x = \beta Q_P(1)$ as indicated in Equation (3).

Updating Assurance Using Bayes' Rule

In the Bayesian paradigm, further audit procedures induce a likelihood function, which updates the auditor's prior assurance profile to a posterior assurance profile. If the prior is $g(x;\alpha,\beta)$ and the likelihood is $g(x;A,B)$, then Bayes' rule (O'Hagan and Forster 2004, 1.9) is:

$$g(x;\alpha',\beta') = \frac{g(x;\alpha,\beta)g(x;A,B)}{\int_0^\infty g(x;\alpha,\beta)g(x;A,B)dx}, \tag{27}$$

where:

$$\alpha' = \alpha + A - 1 \quad \text{and} \quad \beta' = \frac{1}{1/\beta + 1/B}$$

(Raiffa and Schlaifer 2000).

To illustrate, if the prior and likelihood are exponential distributions then from (26) the numerator on the right-hand side of (27) is:

$$\begin{aligned} g(x;1,\beta)g(x;1,B) &= \left(\frac{1}{\beta}e^{-x/\beta}\right)\left(\frac{1}{B}e^{-x/B}\right) = \frac{1}{\beta B}e^{-x(1/\beta+1/B)} = \frac{1}{\beta B}e^{-x/\beta'} = \frac{\beta'}{\beta B}\left(\frac{1}{\beta'}e^{-x/\beta'}\right) \\ &= \frac{\beta'}{\beta B}g(x;1,\beta') = \frac{1}{\beta B}\times\frac{1}{1/\beta+1/B}\times g(x;1,\beta') \\ &= \frac{g(x;1,\beta')}{\beta+B}, \end{aligned}$$

and the denominator is the constant:

$$\int_0^\infty g(x;1,\beta)g(x;1,B)dx = \int_0^\infty \frac{g(x;1,\beta')}{\beta+B}dx = \frac{1}{\beta+B}\int_0^\infty g(x;1,\beta')dx = \frac{1}{\beta+B}.$$

Therefore (27) is:

$$g(x;1,\beta') = \frac{g(x;1,\beta)g(x;1,B)}{\int_0^\infty g(x;1,\beta)g(x;1,B)dx} = \frac{g(x;1,\beta)g(x;1,B)}{1/(\beta+B)},$$

or, as illustrated by Equation (7):

$$\underset{\text{Posterior Density}}{g(x;1,\beta')} = \underset{\text{Prior Density}}{g(x;1,\beta)} \times \underset{\text{Likelihood Function}}{g(x;1,B)} \times \underset{\text{Constant}}{(\beta+B)}.$$

APPENDIX B

OPTIMALITY OF SQUARE-ROOT WEIGHTING

In this appendix we show that if the objective is to minimize component audit costs, then components should be weighted approximately in proportion to the square root of their size as indicated in square-root weighting Formula (24).

We use Equations (17) and (22) to express total variable cost as a function of the weights:

$$TVC = \sum TVC_i = \sum \frac{R_{0.95}Y_i\left(Q_{0.95}(1/w_i) - T/\beta_i\right)}{TR_{0.95}} = \sum \frac{Y_i Q_{0.95}(1/w_i)}{T} - \sum \frac{Y_i}{\beta_i},$$

where Y_i is revenues, total assets, or some other appropriate financial benchmark. Minimizing TVC with respect to $(w_1,\ldots,w_N)$ is a classic constrained optimization problem in which the goal is to find the weights that minimize the objective function:

$$f(w_1, w_2, \ldots, w_N) = \sum Y_j Q_{0.95}(1/w_j),$$

subject to the constraint $\sum w_j = 1$. Following the Lagrange method (Cooper and Steinberg 1970, 120), we construct:

$$F(w_1, w_2, \ldots, w_N, \lambda) = \sum Y_j Q_{0.95}(1/w_j) + \lambda(1 - \sum w_j),$$

differentiate it with respect to the w_i and λ, and set the partial derivatives equal to zero, thus obtaining:

$$F'_{w_i}(w_1, w_2, \ldots, w_N, \lambda) = \frac{-Y_i Q'_{0.95}(1/w_i)}{w_i^2} - \lambda = 0 \quad \text{for} \quad i = 1, 2, \ldots, N,$$

$$F'_{\lambda}(w_1, w_2, \ldots, w_N, \lambda) = 1 - \sum w_j = 0,$$

where $Q'_{0.95}$ is the derivative of $Q_{0.95}$ and F' is the partial derivative of F with respect to the subscripted variable. Solving for w_i gives:

$$w_i = \frac{\sqrt{Y_i Q'_{0.95}(1/w_i)}}{\sum \sqrt{Y_j Q'_{0.95}(1/w_j)}}.$$

It can be shown that the derivatives make no appreciable difference and can be ignored (Stewart 2012), thus leading to weighting Formula (24).

THE ACCOUNTING REVIEW
Vol. 88, No. 2
2013
pp. 739–747

American Accounting Association
DOI: 10.2308/accr-10324

Book Reviews

Stephen A. Zeff, Editor

Editor's note: Two copies of books for review should be sent to the Book Review Editor: Stephen A. Zeff, Rice University, Jesse H. Jones Graduate School of Business, 6100 Main St., Houston, TX 77005. The policy of *The Accounting Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

JENNIFER BLOUIN, *Taxation of Multinational Corporations*, Foundations and Trends® in Accounting (Hanover, MA: now Publishers, Inc., 2012, ISBN 978-1-60198-532-3, vol. 6, no. 1, pp. x, 64).

A Google Scholar search of the title "Taxation of Multinational Corporations" gives approximately 44,500 hits. In conjunction with "Accounting," "Finance," and "Public Economics," the search results are similar. These numbers show that international taxation is a vital research area that deals not only with fiscal borders, but also crosses the borders of academic disciplines.

The monograph *Taxation of Multinational Corporations* by Jennifer Blouin (Wharton School, University of Pennsylvania) is the most recent issue in the series entitled "Foundations and Trends® in Accounting." The objective of this series is to guide academics through the most important literature in a certain field of accounting research by means of high-quality surveys. The objective of this number in the series is to familiarize accounting researchers with current research questions and major publications in international taxation, regardless of whether the papers were published in accounting, finance, or public economics journals. This monograph should prove useful for non-tax accounting researchers, as well as for doctoral students in tax accounting interested in the issues arising in international taxation.

By focusing on the impact of taxation on decisions by U.S. multinationals related to their cross-border activities, the monograph complements the excellent review of the tax literature by Hanlon and Heitzman (2010). Citation of this review article in the monograph would have constituted valuable information for the interested reader.

The monograph is structured in seven chapters. Chapter 1 introduces the reader to various tax-driven problems and decisions faced by multinational corporations (MNCs); among them, location decisions as the most influential decisions from a welfare and tax revenue perspective, as well as repatriation decisions and profit-shifting incentives. The author correctly points to tax avoidance by MNCs that may result from overlapping tax claims by at least two jurisdictions. Although tax avoidance incentives are indeed a major threat to national tax revenues, the efficiency-threatening effect of double taxation could be initially mentioned.

There exist two basic systems of avoiding international double taxation as described in the monograph: a territorial system exempting foreign income in the domestic country, and a worldwide system crediting foreign taxes against the domestic tax liability. Although the majority of OECD countries have implemented territorial systems, most of the empirical research in top-level academic accounting journals is based on data of U.S. MNCs facing a worldwide system. The author openly admits the resulting U.S.-centric view of the monograph (abstract and p. 3). It could be noted that U.S. data are less dominant in related fields of tax research, e.g., in public economics, in analytical tax research, or in second-tier journals. However, given the aims and scope of the monograph series focusing on research published in top-level journals, this limitation seems justified.

Chapter 2 provides an overview of the taxation of U.S. MNCs, focusing on the principle of worldwide taxation, tax deferral, and the foreign tax credit (FTC). Under worldwide taxation, total income from the worldwide operations of a U.S. firm and its branches and subsidiaries is subject to U.S. taxation. Whereas the

U.S. income and foreign income of the U.S. parent company and its branches are immediately taxable, the income of foreign affiliates organized as separate subsidiaries is taxable only upon repatriation, but not upon realization (deferral). Double taxation is avoided by crediting foreign taxes on foreign-source income against the U.S. tax liability. In contrast to most jurisdictions with territorial systems, the U.S. foreign tax credit (FTC) is not limited to the direct credit of foreign taxes paid by the parent company, but also includes the indirect credit of foreign taxes paid by subsidiaries. The FTC limitation on aggregate foreign-source income (the so-called overall limitation) is another element that distinguishes the U.S. from other countries, which typically practice a per-country limitation. Further U.S.-specific characteristics of the FTC are the income-basket limitation and an FTC carryforward and carryback that are denied by many jurisdictions.

All of these principles of U.S. international tax law are clearly presented. Non-tax accounting researchers and non-U.S. tax researchers alike will gain useful insights from this description. Only one paragraph on Subpart F rules (p. 8) requires knowledge of the technical details of U.S. corporate law. Brief historical notes on past legislation, e.g., on controlled foreign corporations (Subpart F) legislation and the foreign tax credit, enhance the understanding of current U.S. tax law. Additional suggestions for further reading on U.S. tax law, although not being top-level research, would have been helpful for non-U.S. researchers.

Chapter 3 is the most detailed part of the monograph. This chapter discusses the impact of taxation on investment and especially on repatriation decisions. Section 3.1 is designed to provide a theory of the impact of taxes on the investment decisions of MNCs. Actually, this section contrasts the concept of capital export neutrality with capital import neutrality and the corresponding systems of worldwide versus territorial taxation. Capital ownership neutrality, as an extension of these concepts, is not mentioned (see Desai and Hines 2003).

The analytical part of Section 3.1 (pp. 16–18) shows how the difference between U.S. and foreign taxes affects the initial capital contribution of foreign subsidiaries, and proves that repatriating and recontributing capital back to the subsidiary ("roundtripping") cannot be optimal. The understanding of this paragraph (p. 17) is complicated by a typo in the equations (Equations (3.4) and (3.6) are identical, but their difference in (3.7) is non-zero). The survey part of Section 3.1 mainly refers to research from the 1980s to the 2000s.

Section 3.2 deals with the repatriation decisions of U.S. MNCs. The author first presents a simple analytical model of repatriation that clearly shows the tax incentives as functions of after-tax rates of return. Thus, it becomes obvious that tax rate differentials determine repatriation policy. It would have been a helpful extension to add that real-world investment and repatriation decisions are essentially multi-period problems under uncertainty with multiple intertwined decision variables. If an MNC has n subsidiaries and a planning horizon of T periods, there are $n\,T$ non-negative repatriation variables, which are all interdependent due to the overall limitation and the carryover rules of the FTC. In addition, $n\ T$ real-valued investment variables have to be taken into account. The reader receives a slight impression of the complexity of real-world tax planning later in Section 3.2.4, where some more details of the FTC, e.g., the differentiation between direct and indirect tax credit, are mentioned.

An overview of the empirical studies on the impact of taxation on repatriation in Section 3.2.2 refutes earlier theoretical results, predicting the irrelevance of repatriation taxes, and shows that repatriations are a decreasing function of the U.S.-foreign tax rate differential. Potential explanations for this effect are varying definitions of taxable income and a non-stationary FTC position. By referring to the separation of permanent and transitory components of repatriation taxes, the author points to a hitherto unresolved research question.

Consequently, much attention is devoted to repatriation decisions following the 2004 American Jobs Creation Act (AJCA), which temporarily reduced the tax cost of repatriations and, thus, represents an example of varying repatriation tax rates. The author derives conditions for optimal repatriation under the AJCA and provides empirical evidence that repatriating firms had substantially higher share repurchases than non-repatriating firms, although repatriating firms were required to invest in approved activities.

Given the usual tax penalty on repatriations to the U.S., the question arises whether mechanisms other than dividends can be used as repatriation devices. This issue is addressed in Section 3.2.5. Tax rate differentials induce an obvious incentive to finance subsidiaries in high-tax countries with debt and to finance subsidiaries in low-tax countries with equity. On page 32, the author states that "the withholding tax rates on the tax-deductible remittances may be substantially higher than the withholding rates on dividends." While withholding tax rates are sometimes higher for royalties than for dividends, it should be noted that many double taxation treaties have lower withholding tax rates on interest income than on dividends, aggravating the above-mentioned financing incentive. The cited literature provides confirming evidence.

The remarks on tax havens in Section 3.3 lead the reader to the income shifting and transfer pricing decisions discussed in Chapter 4, constituting the second main part of the monograph. The introduction and the theoretical part of this chapter clearly describe the tax-induced incentives for income shifting within MNCs by means of adjusting transfer prices for intra-group sales. In the theoretical literature, it has been intensively discussed whether a single transfer price or separate transfer prices should be used for taxation and managerial reporting purposes. (For an integrated view, see Hiemann and Reichelstein [2012].) This important issue is briefly addressed with a single analytical reference. Some empirical evidence would have been helpful.

The overview of the empirical literature on income shifting is very detailed. Even for tax researchers, it may be novel that tariffs rather than taxes are a pivotal determinant of transfer pricing. Although intra-group debt financing is a major device for international profit shifting, this issue is only incidentally mentioned. This is also true for thin capitalization rules, which have been implemented as a fiscal countermeasure against excessive debt financing of subsidiaries in high-tax jurisdictions (p. 39, fn. 4). Against the background of formula apportionment as the current intra-U.S. and proposed intra-EU tax allocation mechanism, it would have been a useful extension to address the problem of high compliance costs associated with transfer pricing documentation requirements.

In Chapter 5, the author highlights the interplay of tax- and non-tax considerations for the location and repatriation decisions of U.S. MNCs. This chapter suggests that the international accounting literature has only recently picked up the interdependence of tax incentives and financial reporting incentives, although at least three academic disciplines have intensively investigated international taxation for decades. However, as far as I know, in jurisdictions with book-tax conformity like Germany, the economic effects of interrelated tax and financial accounting have been explored since the 1980s. Against the background that some German tax privileges were discreetly implemented via modifications of local GAAP, it is an interesting anecdote in accounting standard setting that the convergence of IFRS and U.S. GAAP has been affected by lobbying activities related to repatriation tax obligations (pp. 49–50).

The author provides a lucid explanation for the combined tax and financial reporting effects on repatriation decisions by an example on the Indefinite Reversal Exception (p. 48). Obviously, the (non-) declaration of current earnings as permanently reinvested affects the MNC's repatriation tax expense and, hence, the effective tax rate.

Chapter 6 of the monograph presents recent developments in the taxation of U.S. MNCs. Apart from mentioning Congressional activities to confine tax sheltering, this chapter focuses on the current discussion of moving from a worldwide to a territorial system of corporate taxation.

With only 14 printed lines, Chapter 7 is a very brief conclusion of the monograph. This chapter would have been ideal for suggestions for further research in international taxation.

Evidently, the taxation of multinational corporations is an extremely large and diversified research area, so that a single survey is unable to cover all of the complex aspects arising in this context. Consequently, some research questions are neglected in this monograph; among them, group taxation, the taxation of relocation activities, AMT issues in international taxation, the use of holding companies, compliance costs, and alternative tax allocation mechanisms like formula apportionment. However, this monograph provides a very valuable extension of the survey by Hanlon and Heitzman (2010) with respect to repatriation decisions, transfer pricing, and non-tax considerations of U.S. multinationals.

REFERENCES

Desai, M. A., and J. R. Hines, Jr. 2003. Evaluating international tax reform. *National Tax Journal* 56 (3): 487–502.

Hanlon, M., and S. Heitzman. 2010. A review of tax research. *Journal of Accounting and Economics* 50 (2/3): 127–178.

Hiemann, M., and S. Reichelstein. 2012. Transfer pricing in multinational corporations: An integrated management- and tax perspective. In *Fundamentals of International Transfer Pricing in Law and Economics*, edited by Schön, W., and K. A. Konrad, 3–18. Berlin, Germany: Springer.

RAINER NIEMANN
Professor of Accounting and Taxation
Karl-Franzens-Universität Graz

KENNETH A. MERCHANT and KATHARINA PICK, *Blind Spots, Biases, and Other Pathologies in the Boardroom* (New York, NY: Business Expert Press, 2010, ISBN-13: 978-1-60649-070-9, pp. x, 151).

Shareholders vote for individual corporate directors, and corporate proxy statements provide descriptions of the experience, qualifications, and skills of the individuals being nominated for election to boards of directors. Yet an individual director only has the power to act on the corporation's behalf as one of a body of directors acting as a board.[1] The central theme of this book is that boards of directors are not mere aggregations of individuals, but complex social systems that exhibit tendencies that occur in groups, including common blind spots, biases, and other similar pathologies. This book is intended for board members, executives, and advanced students who want to learn more about board behavior, particularly the negative effects of group tendencies. One goal of the book is to provide insights for board members and regulators who are interested in improving corporate governance. A second goal is to provide insights for shaping expectations about the degree and quality of advice and oversight that boards can provide, given their inherent limitations.

The book unfolds as follows. Chapter 1 describes group tendencies and biases that can cause highly intelligent, energetic, and well-intentioned individuals to collectively misperceive risks and fail to recognize problems because of how they frame decisions, focus on unimportant details, and suppress dissent. Chapter 2 describes how groups can change individual behavior (social loafing and group conformity), Chapter 3 discusses cognitive limitations of groups because of shared information (shared information bias) and shared concerns (pluralistic ignorance), and Chapter 4 describes the tendency of groups to make riskier decisions (group polarization). Chapters 5 and 6 describe how group conformity (groupthink) can be dysfunctional and can entrench behavior that may be inappropriate for certain situations (group habitual routines), Chapter 7 explains why conflict is inherent in groups, Chapter 8 explores the role and misuse of power, and Chapter 9 discusses how groups can cause productivity losses. Finally, Chapter 10 discusses ways to counteract or overcome these tendencies, as well as dynamics to enhance the effectiveness of boards.

Most chapters begin with a boardroom vignette to motivate the descriptions of the group tendencies and biases, their possible effects, and their functioning in real or plausible board situations. The authors acknowledge that the descriptions have not been developed from studies of boards of directors. Instead, the descriptions rely heavily on findings in experimental social psychology, which the authors believe provide "a nice blend of theory and practice, rigor and relevance" (p. x). Yet this reliance on experimental findings is a potential limitation. Experimental psychologists and economists have long acknowledged the artificial nature of the groups, tasks, events, and settings of experimental research. In particular, the tasks used in many experiments about groups may have experimental but not mundane realism.[2] In particular, the experiments are not likely to reflect or capture the critical evaluative judgments and events of relational boards selecting from among alternative courses of action in fulfilling their advisory and oversight roles and responsibilities.

The authors overcome this limitation by relying on experiments that were used to test theories and by using both the theories and findings to develop the descriptions of the group tendencies and their possible effects. They focus on findings that have been replicated many times, using a wide variety of different tasks and settings. In addition, they focus on theories and findings that they believe are highly relevant for boards, based on their experiences in participating in and observing board meetings in companies and in interviewing board members.

The book describes about a dozen tendencies and biases inherent in all groups. Consider two examples. First, "social loafing" is the tendency for individuals to reduce the effort they put into a task when they are working as part of a group as opposed to working alone. A famous experiment by Max Ringelmann measured how hard subjects pulled a rope (Kravitz and Martin 1986). Subjects pulled a rope harder when they did it individually than when they were pulling as part of a team. This result was partially attributable to the difficulty of coordinating group effort as the size of the group increases, but subsequent studies found that this result also

---

[1] See Bainbridge (2002) for a discussion of boards of directors as groups.

[2] See Swieringa and Weick (1982) for a discussion of experimental realism (whether laboratory events are believed, attended to, and taken seriously) and mundane realism (whether laboratory events are similar to real-world events).

was attributable to individuals believing that their individual efforts will not be visible and identifiable to others, will not necessarily impact group performance, and the task has little meaning or significance for them.

Social loafing is used to explain how highly accomplished and respected individuals can appear to be inattentive, disengaged, or disinterested in the performance of their director duties. To minimize social loafing, the authors suggest that boards structure board discussions, encourage work in standing or *ad hoc* committees, and evaluate directors individually.

Second, "groupthink" is a way of thinking where pressure for unanimity overwhelms an individual's motivation to realistically appraise alternative courses of action. Groupthink is caused by social pressure and a tendency to want to avoid confrontation. Studies by Janis (1982) and others demonstrated that individuals often conform to majority viewpoints, not only because they do not trust their own perceptions, but also because they do not want to stand out from the crowd for fear of being ridiculed. Groupthink is more likely when groups are very cohesive, insulated from outside sources of information, have a strong and directive leader, and are under stress.

Groupthink may cause directors to reach a consensus too quickly because of the way social similarities shape their perceptions and decision making, salient norms reinforce consensus-based decision making, they have unqualified support for the CEO, and open dissent is rare. To minimize groupthink, the authors suggest that boards include members who have different backgrounds and viewpoints, seek expert information from advisors and consultants, ensure that a range of alternatives is considered, create a culture in which members are empowered and expected to debate and challenge each other, and include members with relevant expertise.

The authors focus on specific group tendencies and biases and provide advice about how to mitigate potential negative effects. This approach is another potential limitation, because the underlying group processes and dynamics overlap and interact, resulting in conflicting advice. The final chapter focuses on unavoidable board tensions and some practical advice about avoiding or minimizing the effects of the most serious tendencies and biases, including creating constructive dissent, fashioning the right board composition, and using board evaluations to improve board and corporate performance.

This book is relevant and timely for accountants and auditors interested in improving corporate governance and financial reporting. A recent survey found that a third of audit committee members surveyed indicated that they believed that groupthink tendencies influenced their meetings (Audit Committee Institute 2011). A thought paper from the Committee of Sponsoring Organizations of the Treadway Commission (COSO) highlighted the potential effects of groupthink, as well as bias-inducing tendencies of overconfidence, confirmation, anchoring, and availability on board behavior, and how their effects can be mitigated by seeking opposing and disconfirming evidence, questioning expert opinions, and encouraging opposing points of view (KPMG LLP et al. 2012). Schrand and Zechman (2012) suggest that overconfident executives may be more likely to exhibit an optimistic bias and, thus, may be more likely to start down a slippery slope of growing intentional misstatements.

This book is a valuable resource for understanding and improving board judgments. Most discussions of corporate boards focus on best practices for board size, composition, structure, leadership, focus, and evaluations, but fail to recognize that a board is a group and that group judgments, although often better than individual judgments, can be affected by tendencies and biases inherent in all groups, even boards that have embraced best practices. Awareness of these group tendencies and biases, as well as their symptoms, causes, and possible ways to minimize their effects, is an important initial step in improving board judgments and corporate governance.

REFERENCES

Audit Committee Institute (ACI). 2011. Public company audit committee survey—Highlights. Available at: http://www.nacdonline.org/Resources/Article.cfm?ItemNumber=5276

Bainbridge, S. 2002. Why a board? Group decisionmaking in corporate governance. *Vanderbilt Law Review* 55: 1–55.

Janis, I. L. 1982. *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascos*. Boston, MA: Houghton Mifflin.

KPMG LLP, S. M. Glover, and D. F. Prawitt. 2012. Enhancing board oversight: Avoiding judgment traps and biases. Committee of Sponsoring Organizations of the Treadway Commission. Available at: http://www.kpmg.com/cn/en/issuesandinsights/articlespublications/pages/enhancing-board-oversight-201209.aspx

Kravitz, D. A., and B. Martin. 1986. Ringelmann rediscovered: The original article. *Journal of Personality and Social Psychology* 50 (5): 936–941.

Schrand, C. M., and S. Zechman. 2012. Executive overconfidence and the slippery slope to financial misreporting. *Journal of Accounting and Economics* 53: 311–329.

Swieringa, R., and K. Weick. 1982. An assessment of laboratory experiments in accounting. *Journal of Accounting Research* 20 (Supplement): 56–101.

ROBERT J. SWIERINGA[3]
Professor of Accounting
Cornell University

CHRISTOPHER NAPIER and ROSZAINI HANIFFA (editors), *Islamic Accounting* (Cheltenham, Glos, U.K.: Edward Elgar Publishing Limited, 2011, ISBN 978-1-84844-220-7, pp. xx, 740).

What is to be understood by the term "Islamic accounting"? The question has arisen in the context of the development of the Islamic financial services industry (IFSI) in recent decades. The *raison d'être* of IFSI is Islamic religious law, the *Shari'a*, and its interpretations in Islamic commercial jurisprudence, the *Fiqh al Muamalat*, according to which certain forms of transactions and financial instruments that are widely employed in conventional finance, as well as the conducting or financing of activities connected with alcohol, pork, gambling, and armaments, are prohibited. These include (the Arabic terms are in brackets): any form of interest receipts or payments [*riba*], short selling and speculation in general [*maysir*], and vagueness or ambiguity of contractual outcomes [*gharrar*]. Islamic finance, therefore, uses forms of contract based on *Fiqh al Muamalat* (known as the "nominate contracts") and financial instruments which avoid these prohibited elements. The resulting transactions call for specific accounting treatments that may not be indicated within "generally accepted accounting principles" such as U.S. GAAP or the International Accounting Standards Board's (IASB) IASs and IFRSs. For this reason, a specialized accounting standards body, the Accounting and Auditing Organization for Islamic Financial Institutions (AAOIFI), was set up in 1991, and has since issued around 25 financial accounting standards.

In a narrow sense, "Islamic accounting" might, therefore, be understood as accounting as required by AAOIFI's standards, although such a usage is debatable for reasons I will give below. However, for the editors of this weighty collection of 33 papers (most of which have been previously published in journals, and are reproduced from the originals) plus an introductory chapter, the term has a much wider sense or set of senses. In the introductory chapter, written by the editors and entitled "An Islamic Perspective of Accounting," the narrow sense is presented first. The second sense that they mention is associated with notions of social responsibility within a framework of religious ethics: accounting is seen through the lens of a corporate governance (CG) framework that, in contrast to the neo-liberal and secular perspective that characterizes the U.S. and U.K. approach to CG, with its emphasis on the rights of investors and creditors, stresses accountability to God for socially responsible behavior (including transparency and fair dealing, as well as environmental sustainability). While the first, narrow sense of "Islamic accounting" is concerned with technical issues such as recognition, classification, and measurement, as well as disclosure, this second sense is particularly concerned with the latter, as well as with corporate governance more generally. A third sense of the term is suggested by the need for accounting treatments to provide an appropriate basis for determining liability to *zakat*, a type of wealth tax or obligatory (for Muslims) religious levy intended to finance social causes such as the alleviation of poverty. The book also contains papers concerned with the auditing of Islamic financial

[3] Robert Swieringa has been serving on the board of directors of General Electric Company since 2002.

institutions and with the history of accounting in the Islamic world, as well as a paper on management accounting systems in Islamic and conventional banks.

As its title suggests, the general thrust of the book's contents is Islamic particularism in accounting; namely, the thesis that accounting (in the broad sense together with corporate governance), as envisioned and practiced from an Islamic perspective, is qualitatively and not just technically different from conventional accounting. The second sense of the term "Islamic accounting" is, thus, clearly central to this thesis. However, the book contains only a minority (less than 25 percent) of empirical papers, and it is to these that reference must primarily be made in assessing to what extent the thesis is effectively supported.

After the introductory chapter by the editors, the book is divided into six parts: Conceptual Framework for Islamic Accounting (three papers); Accounting Ethics and Social Responsibility (seven papers); Corporate Reporting (nine papers); Accounting Practice and *Zakat* (seven papers); Auditing (three papers); and Islamic History of Accounting (four papers).

In Part I, a particularly significant paper (Chapter 2) is that by Rifaat Ahmed Abdel Karim, who was at the time secretary-general of AAOIFI. The author explains the rationale for the setting up of AAOIFI (under its initial name, the Financial and Accounting Organization for Islamic Banks and Financial Institutions [FAOIBFI]). The context of the financial reporting issues and problems faced by Islamic banks are also well explained in the previous paper (Chapter 1), by Moustafa F. Adbel-Magid.

Dr. Karim states that two options were considered by FAOIBFI in developing objectives of financial accounting (for Islamic financial institutions):

1. Establish objectives based on the principles of Islam and its teachings and then consider these established objectives in relation to contemporary accounting thought; or
2. Start with objectives established in contemporary accounting thought, test them against Islamic *Shari'a*, accept those that are consistent with *Shari'a*, and reject those that are not. (p. 31)

The author tells us that after a lengthy process of discussions involving accounting academics and practitioners, *Shari'a* scholars, Islamic bankers, and officials in central banks, it was agreed that the second approach should be adopted. Both approaches were considered to be in compliance with *Shari'a* precepts, so that there was no reason to reject the second approach. Moreover, a similar approach was adopted in developing the concepts of financial accounting, which comprised the following:

(a) The identification of accounting concepts which have been previously developed by other institutions, which are consistent with the Islamic ideals of accuracy and fairness.
(b) The identification of concepts which are used in traditional financial accounting but which are inconsistent with Islamic *Shari'a*, which were either rejected or sufficiently modified to comply with the *Shari'a*.
(c) The development of those concepts defining certain aspects of financial accounting for Islamic banks that are unique to the Islamic way of transacting business. (p. 32)

An example of the issue mentioned under (c), above, is the recognition of profit under a type of Islamic credit sale transaction known by the name of the "nominate contract" employed as *Murabaha*, which is a sale at cost (which must be disclosed) plus a mark-up or gross profit margin. Payment of the *Murabaha* price is typically made by installments, which raises the question of how the profit element should be recognized. According to IAS 18, the financial element of the profit (or interest) should be recognized *pro rata temporis*, using the "effective interest rate" method, while the nonfinancial element should be recognized when the sale takes place. However, the *Murabaha* mark-up is not interest, and Islamic banks could interpret IAS 18 in various ways, either to recognize all of the mark-up as profit at the time of the sale, or to recognize it as profit at the conclusion of the contract (when the final payment is made), or, indeed, by various methods of recognition *pro rata temporis*. One might suggest that the mark-up could be decomposed into the "pure interest" element and the "pure gross profit" element by reference to prevailing market interest rates, but that would in fact be an example (to quote from (b), above) of a concept "used in traditional financial accounting but . . . inconsistent with Islamic *Shari'a*." AAOIFI, in its financial accounting standard on *Murabaha*, settled on a method of recognition of the entire mark-up as profit *pro rata temporis*.

The author's explanations of the *raison d'être* of AAOIFI and its standards are amplified in another paper by him, which appears as Chapter 12 in Part III of the collection. A paper in Part IV of the collection, which appears as Chapter 24, by Ros Aniza Mohd. Shariff and Abdul Rahim Abdul Rahman, provides further support

to the proposition that accounting practices for Islamic transactions (in this case, Islamic leases) are divergent in the absence of generally accepted and applicable financial reporting standards.

The papers by Dr. Karim present a particularly well-informed view of what may be meant by "Islamic accounting" in the first, narrow sense mentioned above in a financial accounting standard-setting context, namely, accounting based on the *second* of the two options considered by FAOIBFI/AAOIFI. In contrast, most other papers in the collection seem to favor the *first* of the two options. For example, in the following paper (Chapter 3), Roszaini Haniffa (one of the editors) and Mohammad Abdullah Hudaib take the view that "[b]ased on the limitations of conventional Western accounting, the *Shari'a Islami'iah* is proposed as the foundation in building a theoretical framework for IPA [the Islamic perspective of accounting]" (p. 43). Other papers in the collection also seek to link the notion of "the Islamic perspective of accounting" with a more general critical perspective on accounting; for example, in Chapter 17 by Rania Kamla, Sonja Gallhofer and Jim Haslam, which links Islamic principles and accounting for the environment within a critical perspective.

We have, therefore, in this collection a contrast between:

(a) a minimalist version of the first sense of "Islamic accounting" mentioned above; namely, a practical view of the need to improve the quality of financial accounting and reporting of Islamic financial institutions; i.e., not "Islamic accounting" as such, but rather accounting that fairly presents the results of certain transactions that comply with the Islamic *Shari'a* when the application of conventional U.S. or IASB GAAP would not do so (see the *Murabaha* example above); and
(b) the second, broader, and more radical sense of "Islamic accounting" mentioned above; namely, a view of Islamic accounting (or financial accounting, reporting, and corporate governance) as reflecting Islamic religious ethics in such a way as to make these different in principle from conventional financial accounting, reporting, and corporate governance, which are seen as reflecting a secular, materialist, and arguably amoral world view.

Whether this more radical sense of "Islamic accounting" is significantly reflected in practice is an issue on which the empirically based papers in this collection should be able to shed light. It is, therefore, to these that I will now turn. While none of the papers in Part II, Accounting Ethics and Social Responsibility, are empirically based, three of the papers in Part III, Corporate Reporting, are. The same is true of one paper in Part IV, Accounting Practice and *Zakat*. These papers cast light on the extent to which the more radical sense of "Islamic accounting" is significantly reflected in practice. Three other empirically based papers are included in the collection, Chapters 25 and 26 in Part VI and Chapter 29 in Part V, but as they do not bear on the issue identified above, they will not be examined here.

Chapter 16, a paper by Maliah Sulaiman, reports the results of a "laboratory experiment" using students as surrogates for investors in testing whether certain financial reporting methods developed by Baydoun and Willett (1994, 2000), namely, a current value balance sheet (CVBS) and a value added statement (VAS), which the authors argue are more consistent with Islamic principles than a historical cost balance sheet (HCBS) and a statement of profit and loss (PL), would in practice be found to be preferable to, or more important than, the latter. The results indicated that Muslim subjects did not accord significantly greater importance to the CVBS and VAS compared to the HCBS and PL. Nor were the Muslim subjects' responses significantly different from those of non-Muslims. The author concluded that "the suggestion that Muslims *ought* to be provided with financial information of a different character from what is normally disclosed in Western-based accounting systems, seems to have little support" (p. 380; emphasis in the original).

Chapter 18, a paper by Bassam Maali, Peter Casson, and Christopher Napier (the latter being one of the editors), examines the extent to which social reporting by Islamic banks conforms to a "benchmark set of social disclosures ... derived by applying Islamic principles in an *a priori* manner" (p. 417). The authors conclude: "Contrary to our expectations, the empirical findings suggest that social issues are not of major concern for most Islamic banks" (p. 425).

In Chapter 19, a paper by Roszaini Haniffa and Mohammad Hudaib, the authors "attempt to assess the degree of variation of *communicated* ethical identity ... against a benchmark of *ideal* ethical identity" for Islamic banks (p. 431). This variation is measured by what the authors call the *Ethical Identity Index* (EII). The EII is composed of 78 constructs grouped under eight dimensions. The dimensions are: vision and mission statement; board of directors and top management; products; *Zakah*, charity, and benevolent loans; employees; debtors; community; and *Shari'a* supervisory board (internal supervision of *Shari'a* compliance). Using content analysis, the corporate annual reports of seven Islamic banks for the period 2002–2004 were scored on

the basis of the EII. The highest score was 65 percent, while the lowest was only 16 percent, and the three-year means for the other five banks ranged from 28 percent to 48 percent, which, as the authors note (with surprise), "suggest[ed] a large disparity between the *communicated* and the *ideal* ethical identities" (p. 444, italics in the original).

The findings of these empirically based papers strongly suggest that the more radical sense of "Islamic accounting" is an ideal that is present in the minds of a number of academicians (and maybe others), but is not reflected in practice. In other words, the notion of "compliance with the *Shari'a*," which is reflected in the practice of Islamic banks, does not encompass, to any significant degree, "Islamic accounting" in this sense (including financial reporting and corporate governance). Nevertheless, compliance with the *Shari'a* is the *raison d'être* of Islamic banks. Hence, such compliance must apparently be understood as not implying the need to practice "Islamic accounting" in the more radical sense.

How, then, is such compliance to be understood? What seems to be important is the avoidance of prohibited types of transactions such as charging or paying interest or speculation by short selling, and of strictly prohibited business activities, such as those associated with arms manufacture, alcohol, pork, and gambling, and, in particular, the use of the *Shari'a*-compliant nominate contracts provided by the *Fiqh al Muamalat* (Islamic commercial jurisprudence) as a basis for financial transactions and instruments. Apart from avoidance of the prohibited business activities, the use of such contracts as a basis for an Islamic bank's operations constitutes compliance in a narrow, formal (juristic or legalistic) sense, rather than in the broader and more substantive sense of compliance in practice with Islamic ethical precepts, such as those of socially and environmentally responsible business conduct.

Thus, what appears to be the thesis of the editors of this collection of papers, namely, that accounting, in a broad sense encompassing financial reporting and corporate governance, as envisioned *and practiced* from an Islamic perspective is qualitatively (and not just technically) different from conventional accounting, is not, in fact, borne out when the empirically based papers in the collection are considered. Even the technical differences, due to the need for specific rules for the financial reporting of the results of certain *Shari'a*-compliant transactions, barely constitute "Islamic accounting."

Notwithstanding this lack of coherence, the papers in the collection are well written and well presented (although some are quite old), and a number of them are very informative for readers such as academicians and graduate students interested in Islamic finance. These papers include the editors' introduction, the two papers by Rifaat Ahmed Abdel Karim, and the various empirically based papers cited above. I would not recommend the book to practitioners or to undergraduate students.

REFERENCES

Baydoun, N., and R. Willett. 1994. Islamic accounting theory. Paper presented at the AAANZ Annual Conference, Sydney, Australia.

Baydoun, N., and R. Willett. 2000. Islamic corporate reports. *Abacus* 36 (1): 71–90.

SIMON ARCHER
Visiting Professor in Islamic Finance
University of Reading

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal ... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature. Topical areas of interest to the journal include accounting information systems, auditing and assurance services, financial accounting, management accounting, taxation, and all other areas of accounting, broadly defined. The journal is also open to all rigorous research methods.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and editors are responsible for providing constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow *The Chicago Manual of Style* (15th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary*.

Format

1. All manuscripts should be formatted in 12-point font on 8 1/2 × 11″ paper and should be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins should be at least one inch from top, bottom, and sides.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we".
5. A cover page should show the title of the paper, all authors' names, titles and affiliations, email addresses, and any acknowledgments.
6. The American Accounting Association encourages use of gender-neutral language in its publications.
7. Experimental studies using human subjects should contain a footnote affirming that approval has been granted by the institution where the experiment took place.
8. Headings should be arranged so that major headings are centered, bold, and capitalized. Second-level headings should be flush left, bold, and both uppercase and lowercase. Third-level headings should be flush left, bold, italic, and both uppercase and lowercase. Fourth-level headings should be paragraph indent, bold, and lowercase. Headings and subheadings should not be numbered. For example:

A CENTERED, BOLD, ALL CAPITALIZED, FIRST-LEVEL HEADING

A Flush Left, Bold, Uppercase and Lowercase, Second-Level Heading

A Flush Left, Bold, Italic, Uppercase and Lowercase, Third-Level Heading

A paragraph indent, bold, lowercase, fourth-level heading. Text starts . . .

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.
Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: three days; 3 kilometers; 30 years. All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text; in tables and figures, the symbol % is used.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: a cross-sectional equation; re-form. See *Webster's* for correct usage.

Keywords: The abstract must be followed by at least three keywords to assist in indexing the paper and identifying qualified reviewers.

Abstract/Introduction

An Abstract of about 100 words (150 maximum) should be presented on a separate page immediately preceeding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings. The Keywords statement should appear immediately below the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the Introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

Tables and Figures

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.
6. When information is not available, use "NA" capitalized with no slash between.
7. Figures must be prepared in a form suitable for printing.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Within-text citations are made using an author-year format. Cited works must correspond to the list of works listed in the "References" section. Authors should make an effort to include the relevant page numbers in the within-text citations.

1. In the text, works are cited as follows: author's last name and year, without comma, in parentheses. For example: one author, (Berry 2003); two authors, (Fehr and Schmidt 2003); three or more authors, (Dechow et al. 1998); more than one work cited, (Cole and Yakushiji 1984; Dechow et al. 1995; Levitt 1998); with two works by the same author(s), (Nelson 2003, 2005).
2. Unless confusion would result, do not use "p." or "pp." before page numbers. For example, (Dechow and Dichev 2002, 41–42).
3. When the reference list contains two or more works by the same author (as the only author or first of two or more authors) published in the same year, the suffix a, b, etc., is appended to the date in the

within-text citations and in the "References" section. For example, (Johansson 2004a, 2004b, 2004c; Baiman and Rajan 2002a, 2002b; Dhaliwal et al. 2005a, 2005b).

4. When the author's name is mentioned in the text, it need not be repeated in the citation. For example: "Cohen et al. (2005) provide ..."
5. Citations to institutional works should use acronyms or short titles where practicable. For example: (NCFFR, The Treadway Commission 1987).
6. If the paper refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a "References" section that contains only those works cited within the text. Each entry should contain all information necessary for unambiguous identification of the published work. Use the following formats (which follow *The Chicago Manual of Style*):

1. Arrange citations in alphabetical order according to the surname of the first author or the name of the institution or body responsible for the published work.
2. Use authors' initials instead of proper names.
3. For two or more authors, separate authors with a comma, including a comma before "and" (Dechow, P. M., R. Sloan, and A. Sweeney).
4. Date of publication follows the author's (authors') name(s).
5. Titles of journals or newspapers should not be abbreviated.

Sample entries are as follows:

Baiman, S., and M. Rajan. 2002a. The role of information and opportunism in the choice of buyer-supplier relationships. *Journal of Accounting Research* 40 (2): 247–278.

Baiman, S., and M. Rajan. 2002b. Incentive issues in inter-firm relationships. *Accounting, Organizations and Society* 27 (3): 213–238.

Berry, R. 2003. Testimony before the Senate Committee on Homeland Security and Governmental Affairs Permanent Subcommittee on Investigations. November 18. Available at: http://hsgac.senate.gov/files/111803berry.pdf.

Cohen, D., A. Dey, and T. Lys. 2005. *The Sarbanes Oxley Act of 2002: Implications for Compensation Structure and Risk-Taking Incentives of CEOs*. Working paper, New York University, University of Chicago, and Northwestern University.

Cole, R., and T. Yakushiji, eds. 1984. *The American and Japanese Auto Industries in Transition.* Ann Arbor, MI: University of Michigan.

Dechow, P. M., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

Dechow, P. M., S. P. Kothari, and R. L. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25: 133–168.

Dechow, P. M., and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors, *The Accounting Review* 77 (Supplement): 35–59.

Dhaliwal, D., Erickson, and O. Li. 2005a. Shareholder income taxes and the relation between earnings and returns. *Contemporary Accounting Research* 22: 587–616.

Dhaliwal, D., L. Krull, O. Li, and W. Moser. 2005b. Dividend taxes and implied cost of equity capital. *Journal of Accounting Research* 43: 675–708.

Easton, P. 2003. Discussion of: The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies* 8: 175–183.

Engel, E., R. Hayes, and X. Wang. 2007. The Sarbanes-Oxley Act and firms' going-private decisions. *Journal of Accounting and Economics* (forthcoming).

Fehr, E., and K. Schmidt. 2003. A theory of fairness, competition, and cooperation. In *Advances in Behavioral Economics*, edited by C. Camerer, G. Loewenstein, and M. Rabin, 271–296. New York, NY: Princeton University Press.

Financial Accounting Standards Board (FASB). 2006. *Accounting for Uncertainty in Income Taxes, and Interpretation of FASB Statement No. 109*. FASB Interpretation No. 48. Financial Accounting Series. Norwalk, CT: FASB.

Janis, I. L. 1982. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*. Boston, MA: Houghton Mifflin.

Levitt, A. 1998. The numbers game. Speech delivered at New York University, Center for Law and Business, September 28.

Maggi, G. 1999. The value of commitment with imperfect observability and private information. *RAND Journal of Economics (*Winter) 30: 555–574.

National Commission on Fraudulent Reporting (the Treadway Commission). 1987. *Report of the National Commission on Fraudulent Financial Reporting*. Washington, D.C.: NCFFR.

Nelson, M. W. 2003. Behavioral evidence on the effects of principles- and rules-based standards. *Accounting Horizons* 17 (1): 91–104.

Nelson, M. W. 2005. A review of experimental and archival conflicts-of-interest research in auditing. In *Conflicts of Interest: Challenges and Solutions in Business, Law, Medicine, and Public Policy*, edited by D. A. Moore, D. M. Cain, G. Loewenstein, and M. H. Bazerman. Cambridge, U.K.: Cambridge University Press.

Rigdon, E. E., R. E. Schumacker, and W. Wothke. 1998. A comparative review of interaction and nonlinear modeling. In *Interaction and Nonlinear Effects in Structural Equation Modeling*, edited by R. E. Schumacker, and G. A. Marcoulides, 1–16. Mahwah, NJ: Erlbaum Associates.

Schultz, E., and T. Francis. 2002. Companies profit on workers' deaths through "dead peasants" insurance. *Wall Street Journal* (April 19): 1.

Securities and Exchange Commission (SEC). 2002. *Certification of Disclosure in Companies' Quarterly and Annual Reports*. Release Nos. 33-8124, 34-46427. Washington, D.C.: SEC.

U.S. House of Representatives. 2002. The Sarbanes-Oxley Act of 2002. Public Law 107-204 [H. R. 3763]. Washington, D.C.: Government Printing Office.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be inserted using the "footnote" or "endnote" feature of the word processing software which will automatically number the footnotes throughout the manuscript with superscript Arabic numerals.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state upon submission that the work is not submitted or published elsewhere.
2. For manuscripts reporting on field surveys or experiments: If the additional documentation (e.g., questionnaire, case, interview schedule) is sent as a separate file, then all information that might identify the author(s) must be deleted from the instrument.
3. New manuscripts must be submitted through the Manuscript Submission and Peer Review System, located at http://accr.allentrack.net. The site contains detailed instructions regarding the preparation of files for submission. Revisions of manuscripts originally submitted before June 1, 2011 should be sent via email to the senior editor, John Harry Evans III, at tar@katz.pitt.edu. Please submit separate files for (1) the manuscript's title page with identifying information (not forwarded to reviewers), (2) the manuscript with title page and all other identifying information removed, and (3) any necessary supplemental files, such as experimental instructions and/or response memoranda on invited revisions.
4. A nonrefundable submission fee is required unless explicitly waived by the editor for invited revisions of previous submissions. The submission fee in U.S. funds is $200.00 for members and $400.00 for nonmembers of the AAA payable by credit card (VISA or MasterCard only). The payment form is available online at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If you are unable to pay by credit card or have any questions, please contact the AAA Member Services Team at (941) 921-7747 or info@aaahq.org.
5. Revisions must be submitted within 12 months from the decision letter inviting a revision.

COMMENTS

The Accounting Review welcomes submissions of comments on previous *TAR* articles. Comments on articles previously published in *The Accounting Review* will generally be reviewed by two reviewers, usually including an author of the original article (to assist the editor in evaluating whether the submitted comment represents the prior article accurately) and an additional reviewer who is independent of the original article. If a comment is accepted for publication, the original author will generally be invited to reply. All other editorial requirements, as enumerated above, apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The AAA's Executive Committee policy (originally adopted in 1989, and amended in 2009) is that the objective of the Association-wide journals (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. To fulfill this objective, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles.

MARCH 2013 PLACEMENT ADS

The deadline for free position ads to be included in this section of *The Accounting Review* is two months prior to the desired publication in the January, March, May, July, September, or November issues. Position ads, which are free with the purchase of a job posting in the AAA Career Center, should provide all relevant information about the available positions and must include contact information or application instructions for interested candidates. For more information on how to purchase a job posting in the Career Center, or for up-to-date, detailed information about the placement listings in this issue, please go to the AAA website at http://aaahq.org and click on "Career Center," or call our office at 941-921-7747.

THE UNIVERSITY OF IOWA, The Department of Accounting, Tippie College of Business, University of Iowa, Iowa City, Iowa, invites applicants for a tenure-track appointment(s) beginning Fall 2013. We will consider candidates at all ranks. Required qualifications include a demonstrated interest in and capacity to do research publishable in top-tier accounting journals, a high level of teaching competence, and an earned doctorate (entry-level candidates can be near completion). We offer a competitive compensation package. We will screen applicants at all ranks on an ongoing basis. To apply, please log on to our website, http://jobs.uiowa.edu/jobSearch/faculty/ and reference Requisition #61273. You will be asked to upload a résumé, representative research output, evidence of teaching capabilities (if available); and to provide the names for at least three letters of recommendation. The University of Iowa is an Affirmative Action/Equal Opportunity Employer. Women, minorities, and individuals with disabilities are encouraged to apply.

DONGBEI UNIVERSITY OF FINANCE AND ECONOMICS (DUFE), China seeks applicants for Assistant Professor, Associate Professor, Research Fellow, and Lecturer. Job duties for a successful applicant will include research, teaching (which is not required if applying for a researcher position), and service to the School and University as full-time staff member. BENEFITS: The annual salary will be RMB 300,000 or above, plus health insurance, housing accumulation funds, temporary housing (or housing allowances), traveling expenses for international conference, research grant, and Ph.D. allowance. Rank, extra awards, and funds will be granted depending on research experience and performance. Additionally, private office and necessary research resources are available. Candidates must have a completed Ph.D. degree in accounting or corporate finance from a prestigious overseas university and concentrate on the research of accounting, auditing, or corporate finance. Contact Qi Yuan at: qiyuan@dufe.edu.cn.

UNIVERSITY OF PITTSBURGH, The Joseph M. Katz Graduate School of Business and College of Business Administration seeks applications for an anticipated tenure-track Assistant Professor position in the accounting area starting Fall 2013 (contingent on budgetary approval and upon authorization to work in the United States). The appointment requires a strong commitment to excellence in scholarly research and teaching. All areas of accounting are appropriate, but preference will be given to experimental researchers. Applications must be submitted electronically to: http://www.katz.pitt.edu/facultyhiring/accounting-ts.php. Materials submitted must include a letter of interest, *vitae*/resume, sample research paper(s), evidence of teaching effectiveness, and three letters of reference. Application deadline is January 15, 2013. Please apply online to the website listed above, but questions about the position can be directed to accfacultysearch@katz.pitt.edu. The University of Pittsburgh is an Affirmative Action, Equal Opportunity Employer and values equality of opportunity, human dignity, and diversity.

BONAVENTURE UNIVERSITY, The Department of Accounting at invites applications for a tenure-track position at the Assistant or Associate Professor level. The successful candidate will teach undergraduate and graduate courses, predominantly in the area of Financial Accounting and/or Audit. The candidate will also be expected to undertake scholarly research in a collaborative environment that welcomes inter-disciplinary and practice-oriented scholarship, as well as discipline-based and theoretical work. The position appointment is expected to begin with the Fall 2013 semester, contingent upon funding. Applicants must meet the following required qualifications: (1) earned doctorate in accounting or related field by the time of appointment; (2) demonstrated excellence in teaching and willingness to work closely with students in curricular and co-curricular activities, including service, professional and/or research activities; and (3) an established or emerging research program. Accounting certification and prior professional accounting experience are preferred. St. Bonaventure University is the nation's premier Franciscan university. The School of Business, AACSB-accredited, will be moving into its new state-of-the-art Swan Business Center home in Fall 2013. Accounting faculty, students, and programs also benefit from the endowed McQuade Center for Accounting Excellence. The Department of Accounting offers two program tracks: a 120-hour B.B.A. in Accounting and a 150-hour B.B.A./M.B.A. program in Accounting registered for CPA Exam qualifications. Our B.B.A./M.B.A. graduates are highly recruited by Big 4 CPA firms and other large employers. Applications may be submitted online at: accountingsearch@sbu.edu. For additional information please contact the department chair, Dr. Susan B. Anders at sanders@sbu.edu or (716) 375-2063. Dr. Anders will be in attendance at the AAA Annual Meeting in Washington, DC. Review of applications will begin immediately and continue until the position is filled. The University, School, and Department have a strong commitment to achieving diversity among faculty and staff, and encourage applications from members of underrepresented groups.

UNIVERSITY OF NORTHERN IOWA has available a tenure-track position as Assistant, Associate, or Full Professor of Accounting available August 2013. The ideal candidate will have an interest in teaching taxation courses. Required qualifications include a Ph.D./D.B.A. in Accounting from an AACSB-accredited institution (ABDs near degree completion will be considered) or an L.L.M. in Taxation (with an accounting background) from an ABA-accredited institution; potential for excellence in teaching; potential to produce publishable research; and evidence of strong communication and interpersonal skills. Professional certification preferred. Complete applications received by February 1, 2013 will be given full consideration. Candidates must apply online. To apply, visit jobs.uni.edu/faculty/view/50132. For more information please contact Dr. Dennis Schmidt, Chair, Accounting Search Committee. Phone: (319) 273-6380; Email: dennis.schmidt@uni.edu. The University is an Equal Opportunity Employer with a comprehensive plan for Affirmative Action.

NEW YORK INSTITUTE OF TECHNOLOGY (NYIT), School of Management seeks qualified candidates for two full-time teaching positions (tenure-track Assistant Professors) in Accounting (specialty in Auditing and Financial Accounting or Tax) program at its New York campus locations. The applicant must hold a Ph.D. (or be ABD) from an AACSB-accredited institution and have demonstrated potential in scholarship within the area of expertise, and a demonstrated commitment to service to students and the local community. Familiarity with discipline-specific software and technology is highly encouraged. Salary is commensurate with qualifications. Selected candidates will be expected to teach courses in accounting and related areas effectively in the Bachelor's and Master's programs; conduct research, publish in peer-reviewed journals, advise/mentor students and participate in the usual faculty duties, including services for students and local community. School of Management offers degree programs at the bachelor's (B.S.) and master's (M.B.A.) level. More information about the school and current faculty can be found at http://nyit.edu/management. Interested candidates should send their *cv*, references, etc., electronically to Dr. J. K. Yun (Email: jyun04@nyit.edu; Phone: 516-686-1173). NYIT hires on the basis of merit and is committed to employment equity.

UNIVERSITY OF NORTH TEXAS, Department of Accounting (http://www.cob.unt.edu/acct) in Denton, Texas (DFW Metroplex) invites applications for a tenure-track faculty position at the Assistant/Associate Professor rank. The discipline targeted for this search is open with preference for financial accounting, accounting information systems, or taxation. The Department of Accounting at UNT offers a full range of academic degree programs including Bachelor of Accounting (B.B.A), professional 150 hour program (B.S./M.S.) in Taxation or Accounting, M.S. in Taxation or Accounting, and an Accounting doctoral program. Our programs are among the most innovative in the country, our faculty exhibit expertise in a host of accounting disciplines, and the department is separately AACSB-accredited. Candidates must hold a doctorate in accounting (or equivalent) and demonstrate excellence in teaching, a strong publication record (for Associate Professor) or a strong publication potential (for Assistant Professor). Candidates for the Assistant Professor must have completed all doctoral degree requirements by start of employment and exhibit potential to achieve the highest standards of scholarship and teaching excellence. CPA, CMA, CIA or equivalent certification is preferred, but not required. The Department of Accounting is committed to academic excellence and diversity within our faculty, staff, and student body. Apply at: https://facultyjobs.unt.edu/

THE GEORGE WASHINGTON UNIVERSITY, School of Business, The Department of Accountancy of invites applications for a tenure-track position at the Assistant Professor level to begin Fall 2013. Basic Qualifications: Applicants must have a Ph.D. in accounting or a related field. ABD candidates will be considered but must complete all doctoral degree requirements by August 1, 2013. Applicants must demonstrate research excellence/strong potential as indicated by refereed publications in top journals or works in progress and teaching excellence as evidenced by letters of reference and teaching evaluations. Application Procedure: To be considered, please complete the online faculty application at: http://www.gwu.jobs/postings/12572 and upload a cover letter, current *curriculum vitae*, up to three working papers and summary teaching evaluations, if available. Three letters of reference must also be submitted, addressed to the Chair, Department of Accountancy, The George Washington University School of Business, Funger Hall 601, 2201 G Street NW, Washington, DC 20052 (email: accyjob@gwu.edu) Only complete applications will be considered. Review of applications will begin on January 14, 2013 and will continue until the position is filled. The George Washington University is an Equal Opportunity/Affirmative Action Employer. The University and the GW School of Business seek to attract an active, culturally and academically diverse faculty of the highest caliber.

UNIVERSITY OF INDIANAPOLIS, Accounting School of Business, The School of Business at the University of Indianapolis seeks applicants for an Assistant or Associate Professor of Accounting. This is a full-time, nine-month, tenure-track position with a start date of August 2013. The successful candidate will participate in teaching, service, and scholarship as it relates to the field of accounting. The ideal candidate should have a strong record of teaching excellence at the undergraduate level for both accounting and non-accounting majors. The ability to teach introductory financial and managerial accounting, cost accounting, and accounting information systems is required. Teaching some evening sections is also required. Qualifications: Candidates should have an earned doctorate in accounting or closely related business field for this tenure-track position. ABD candidates may apply. Professionally qualified candidates with an M.B.A./CPA may be considered for a non-tenure-track appointment. Candidates must have demonstrated proficiency in teaching and knowledge of the accounting field. Applied scholarship and active participation in the discipline are plusses. Apply electronically at: https://jobs.uindy.edu. Review of applications will begin immediately and continue until the position is filled. The University of Indianapolis is an Affirmative Action/Equal Opportunity employer and encourages applications from women and minorities.

UNIVERSITY OF MASSACHUSETTS LOWELL, Robert J. Manning School of Business (MSB) invites applications for multiple tenure-track positions in Accounting at the rank of Associate/Assistant Professor. The MSB is AACSB-accredited. A new M.S.A. program was launched in September 2012. In addition, the MSB expects to add an accounting track to its new Ph.D. in Business Administration in fall 2013 and to occupy a new state-of-the-art instructional facility within the next three years. Lowell, with an ethnically diverse population of approximately 100,000, is located about 30 miles northwest of Boston. Requirements: An earned Ph.D. or D.B.A. in accounting or a related discipline by the starting date (September 2013 or later). A demonstrated record of, or at the assistant level a strong potential for, high-quality research. A demonstrated record of high-quality instruction in some combination of Financial/ Managerial Accounting, AIS, Auditing, or Taxes, at both the graduate and undergraduate levels. Excellent interpersonal skills and the ability to speak and write English clearly. The ability to work with a diverse student and faculty population. Note: Strong consideration given to candidates with public accounting or corporate experience, preferably with CPA, CMA, etc. Applying: Applicants should apply online at: http://jobs.uml.edu Attach a (1) a statement of professional interests and goals, (2) *cv*, (3) samples of research, (4) teaching philosophy and evidence of teaching effectiveness. Applications accepted until the posting is closed.

WASHINGTON & LEE UNIVERSITY, Department of Accounting invites applications for a tenure-track position at the Assistant or Associate level beginning in Fall 2013. Washington & Lee is a small, private, highly selective university located in the scenic Shenandoah Valley. The Williams School of Commerce, Economics, and Politics has been AACSB-accredited since 1927. Job Qualifications: Applicants should have a doctorate or be close to completion and have teaching interests in auditing. In addition to requiring high-quality teaching, applicants are expected to conduct research publishable in peer-reviewed journals. To support research, the school subscribes to several popular databases available through WRDS. Application Procedure: Candidates should electronically submit a cover letter, *curriculum vitae*, evidence of teaching effectiveness and scholarly work, and three letters of reference to: http://jobs.wlu.edu/postings/1198. Specific questions about the position should be sent via email to Dr. Afshad Irani (irania@wlu.edu) who is heading the search. Application review will begin immediately. Washington & Lee University is an Equal Opportunity/Affirmative Action Employer.

FROSTBURG STATE UNIVERSITY, College of Business seeks applications for a full-time, tenure-track Instructor/Assistant Professor of Accounting for its Department of Accounting. The primary teaching area is accounting. Start date is Fall 2013. Salary is competitive and commensurate to experience and qualifications and includes USM benefits package. Responsibilities: Teaching a variety of accounting courses in areas of Financial Accounting, Managerial Accounting or others at the undergraduate and graduate level including online delivery. Occasional travel to the USM Hagerstown Educational Center may be necessary. Peer-reviewed scholarly work and service are required. Frostburg State University's College of Business is AACSB-accredited and values teaching, scholarly contribution, and service. Minimum Qualifications: ABD candidates with a firm completion date will be considered for appointment at the rank of Instructor. Previous teaching experience and evidence of peer-reviewed scholarly work are required. Preferred Qualifications: Earned doctorate in Accounting or related field. Full-time Accounting experience and professional certifications (CPA, CMA) are desired. Application review to begin immediately; applications received until March 8, 2013. To apply, email a letter of interest, resume, and the names, addresses, email addresses, and telephone numbers of three professional references, to: humanresources@frostburg.edu. Include "Instructor/Assistant Professor of Accounting" (Position #13-010539) in the subject line. Frostburg State University welcomes and encourages women and minorities to apply and seeks to recruit and retain a diverse workforce. FSU is an Equal Opportunity Employer. Appropriate auxiliary aids and services for qualified individuals with disability will be provided upon request. Please notify us in advance. See http://www.frostburg.edu

Request for Proposals for Research on the Middle Market

National Center for the Middle Market (NCMM)
Fisher College of Business
The Ohio State University
Columbus, OH USA

Deadline for Submission: May 17, 2013
Funding Decisions: June 16, 2013
Research Period: June, 2013 – July, 2014

Topics: All topics dealing with Middle Market firms, particularly those with impact on growth.

Proposal: The Center invites proposals from individual researchers or teams of researchers. Proposals should be limited to 5 pages plus necessary appendices. For full details on the proposal and obligations of researchers, please visit: http://www.middlemarketcenter.org/research-grant-proposal

Funding: Each individual researcher can expect support of up to 10% of 9-month base salary, plus limited research related expenses.

Center: The NCMM is dedicated to research, corporate outreach and student-related activities concerning the Middle Market, which some have defined as firms with revenues between 10 million and 1 billion annually, an important but understudied segment of the economy. Please visit the Center at: http://www.middlemarketcenter.org/

For questions please email Professor Anil K. Makhija, Academic Director, at middlemarketcenter@fisher.osu.edu

THE ACCOUNTING REVIEW
ISSN 0001-4826

PERIODICALS POSTAGE PAID
SARASOTA, FLORIDA
AND ADDITIONAL MAILING OFFICES

POSTMASTER
Send Address Changes To:

AMERICAN ACCOUNTING ASSOCIATION
5717 BESSIE DRIVE
SARASOTA, FL 34233